The Alberta Journal of Educational Research

Vol. V, No. 3 *September, 1959*

THE COMMITTEE ON EDUCATIONAL RESEARCH
Faculty of Education
University of Alberta

ACKNOWLEDGMENT

This publication was made possible by funds granted by Carnegie Corporation of New York, That Corporation is not, however, the author, owner, publisher, or proprietor of this publication, and is not to be understood as approving by virtue of its grant any of the statements made or views expressed therein.

The Alberta Journal of Educational Research

THE COMMITTEE ON EDUCATIONAL RESEARCH

Faculty of Education
University of Alberta

VOL. V, No. 3 SEPTEMBER, 1959

Editor: H. E. Smith

CONTENTS

PAGE

A Visit to Centres of Educational Research in the Western Universities of the United States 153
Charles C. Anderson

Problem Solving and Achievement 166
Marion D. Jenkinson and Dorothy M. Lampard

Fifty Years after Alfred Binet 173
B. E. Walker

A Descriptive Survey of the Extra-curricular Programs of the Composite High Schools of Alberta 176
Arthur Kratzman and A. W. Reeves

A Comparison of the Performance of Selected Standardized Tests to that of the Alberta Grade XII Departmental Examination of a Select Group of University of Alberta Freshmen 180
D. B. Black

Promotion Vs. Nonpromotion: II. The Edmonton Study 191
Walter H. Worth

Explanation and the Neo-Marxist Theory of Mental Testing 204
C. C. Anderson

Digitized by the Internet Archive
in 2025

https://archive.org/details/alberta-journal-educational-research_1959-09_5_3

A VISIT TO CENTRES OF EDUCATIONAL RESEARCH IN THE WESTERN UNIVERSITIES OF THE UNITED STATES[1]

CHARLES C. ANDERSON
Faculty of Education
University of Alberta

In December 1958 I received a grant from the Carnegie Corporation of New York which enabled me to visit centres of educational research in the western universities of the United States in April and May of 1959. Although my primary purpose was to examine current American research, it rapidly became obvious that the research was only one aspect (albeit a very important one), of a larger theme: the quality and justification of the professional study of education at a university as distinct from a training college level. This change in the focus of my enquiry arose from two sources: firstly the sweeping criticism to which Faculties and Schools of Education were subjected, and secondly the relative poverty of the intellectual training they purveyed.

With regard to the former, it was especially noteworthy that everyone, from the lay and intelligent member of the public to the academic, made the same criticism and had the same remedy. The component parts of the critical attitude can be disentangled fairly easily—the main aim of education is cognitive or intellectual training and American teachers have failed to realise this aim, especially with the gifted students; this neglect is to be ascribed to the policies of the Progressive Education Movement which, under the aegis of Kilpatrick, stressed so heavily 'life adjustment' as the main aim of education and the importance of methodology as against 'content' courses; the source of the trouble is the professional educator who accepted uncritically the Kilpatrick doctrine and disseminated it among his students. This attitude is pervasive and powerful and animates the criticism of the professional educator in America.

The most dangerous source of criticism is the popular but intelligent paper in distinguished general periodicals. For example in 1958 *The Atlantic* published two papers, one in April by Lydia Stout entitled "The Certification Racket: Florida and Elsewhere", and one in September by Mortimer Smith, "How to Teach the California Child: Notes from the Never-Never Land". Stout argues that American teaching is being strangled because of the emphasis

[1]In this paper I have listed in References only the relatively technical papers and books and I have paginated the quotations only from them. Quotations from popular articles in newspapers and magazines do not appear in References and are left unpaginated.

laid by professional educators on education courses (these she equates with methodology or what she terms 'how to' courses) as distinct from 'content' courses in arts and science. She terms the former 'Mickey Mouse courses' and argues that, because they are so lacking in intellectual content, they can be passed easily by failures in any other university department. Stout constantly harps on this distinction between 'method' and 'content' courses:

> Even with a doctor's degree in education[1], a teacher may have fewer course credits in subject than are required for a bachelor's degree in liberal arts.

She cites the opinion of the Dean of the College of Education in the University of Florida that

> If you know enough about how to teach, you can always pick up enough facts to teach a high school class.

Teachers are receiving an overdose of pedagogy. Smith's paper stressed this same point with evidence from California.

Implicit in both papers is the unacceptable assumption that the only course which should be offered in a School of Education is that in methodology. However other critics have generously enlarged the scope of educational studies and have sunk their barbs in other, possibly equally inviting, targets. For example Esther B. Wagner, in 'Eagle Day' (Harper's, March 1959), a short story of some quality dealing with the predicament of an eastern teacher in California, reflects on the poverty of her heroine's education:

> Marina's education had nothing to do with anything which the great teacher's certification bodies could recognise as Education. She had a bachelor's degree, with distinction in English from Madson college . . . She possessed an M.A. in English from a massive metropolitan university. All this, in the mysterious world of American public-school teaching, entitled her to a Provisional, or Substandard, Credential, making it barely legal for her to teach, on the condition that she pursue Education in the form of various courses in classroom management (students get sleepy after lunch unless you open the window), adolescent psychology (adolescents are likely to have problems with their skin and their parents), and testing (California Mental Maturity, True-False, Establishment of Norms and Medians).

As one might imagine, Guidance is swatted rather handsomely.

> She attended a course in Guidance Techniques and sent the Professor's mimeographed materials back to her friends at home with *sic* written in her own hand at regular intervals; the materials provided her and her friends with many moments of harmless pleasure.

Indeed most of professional education is a mockery, run by the half-educated for the half-educated.

Another example of the same sort of academic hostility, this time in a large university on the west coast, took the form of an overwhelming support by representatives of all departments of the

[1]In fairness, the same point might be made against the degree of doctor of philosophy in education.

university for a resolution stating that 'methods' courses should not be compulsory for students intending to teach at the higher levels of secondary education. The fact that this criticism is directed at the methodologist does not completely exonerate the educational psychologist and the educational philosopher who are regarded as something less than a blot but yet still a smudge on the academic escutcheon. For example in the September issue of *The Atlantic* a letter was published from an academic member of a state college in Washington which ended with the following comment:

> There is no available evidence that courses in educational philosophy, educational psychology, or similar ones will help to produce better teachers.

This sentiment would be echoed by many members of most of the universities I visited.

Occasionally a teacher or headmaster joins the attack with exactly the same sort of artillery. On April 30, 1959, I attended at Los Angeles a conference organised by the Los Angeles branch of the Educational Testing Service of Princeton, New Jersey. Star of that conference was R. E. K. Rourke who talked brilliantly and at some length on "Recent Curriculum Developments in Mathematics". Implicit but nevertheless persistent in his speech was the assumption that virtually all students in the high school were gifted to the extent that they could attend university and graduate comfortably. Explicit in his speech was the assertion that gifted children required no teaching, that in fact the best thing the relatively duller teacher could do was to get out of the youth's way. As an appendix to this latter point the speaker poured scorn on 'how to teach' courses taught in Schools of Education and stated that students were not being challenged enough by the non-academic courses they were encouraged to take. And Paul Cruikshank, headmaster of the Taft School in Connecticut, writing in *The Atlantic* for July 1958, reinforces Rourke's last statement with an anecdote about a boy with a great deal of scholastic potential but no scholastic training who tried in vain to enter the Taft School. In reply to the question of where the blame be for all this, Cruickshank gives the following anwer:

> The educationists, with their absorption in 'life adjustment' courses at the expense of courses requiring mental discipline and plain hard work, have contributed to producing a generation many of whose members are non-intellectual, regarding education as an unfortunate experience to be passed through as painlessly as possible.

To some extent this criticism is unjust because Schools of Education operate under certain disadvantages which are unique to them. The lecture-load of the education instructor is far higher than that of his university colleague; he is plagued with committee work (the Americans, no less than the Canadians, feel they have solved something by appointing a committee); the supervision of practice-teaching is necessary, onerous and yet time-consuming, and does not re-

dound as research does to the credit of those taking part; and the the youngest instructor in education is usually at least ten years older than a university colleague of similar rank, with correspondingly less energy for the chores of research.. It happens not infrequently that the most experienced teacher who has the most to offer students as a professional educator does not appear to advantage in either publications or research. Indeed it might be argued very plausibly that, as Faculties of Education are organised, two of their functions must conflict: that of training graduate students and that of training good teachers.

While all this should be granted, nevertheless the fact remains that Schools of Education do merit to quite an extent the criticisms so freely thrown at them. The shabbier methodologist who sticks closely to a textbook which has probably been written by an undistinguished but opinionative hand, will simply have to waken up his (or her) ideas and subject them to proper experimental validation. Some day an experiment must be conducted to determine the relative teaching proficiency of two groups equated at least in terms of academic qualifications: the experimental group placed in schools without any teacher training and the control group with teacher training. When I was visiting the University of Southern California, I came across a piece of research like this but (so it seemed to me) certain difficulties in the execution of the experimental design made an adequate decision on the merits of teacher training quite impossible. In the University of Colorado, some methodologists were attached not to the School of Education but to departments in their 'content' fields—the English methodologist belonged to the department of English, and so on. This is a unique solution which I discovered nowhere else.

Further I discovered that most educational psychologists, particularly those who were relatively eminent, looked with disfavour on the rash of courses which has sprung up of recent years in the field of Guidance or Counselling. Most of these seemed to be an anaemic substitute, suitably simplified for the poorer sort of student haunting Schools of Education, for more recognisable disciplines such as remedial education, developmental psychology, the study of personality, clinical psychology and psychometrics. In one large western university, this is dished up to students in courses variously labelled: Pupil Personnel, Counselling and Guidance; Student Personnel Work: Seminar; Student Personnel and Counselling Psychology; Environmental Factors in Counselling Adjustments; Group Guidance; Individual Appraisal in Counselling; Advanced Counselling Theory; and Field Work in Student Personnel and Counselling Psychology. The more I travelled the more I formulated the general axiom that there was an inverse relation between

the quality of a School of Education and the number of Guidance courses it offered.

The villain of the piece in all this is neither Dewey nor Kilpatrick but the American culture. The American still is solid in his belief that teaching is a woman's game and only the very dedicated or intellectually inferior male enters the profession. This is a statistical fact. As Samuel Wiener points out in *The Atlantic* for July 1958, undergraduate and graduate students in education are intellectually the poorest in American universities. He labels them 'the dunces of the universities'. Only when this state of things disappears will the more justifiable and respectable education courses supervene.

The moral of all this for Canadian educators, who have a much larger share of the pool of university ability than their American counterparts, is this: they should provide learning experience which present an intellectual challenge to students at least equal to that presented by other university departments. This can be done partly by increasing the quantity and quality of educational research and partly by forging strong links between courses in education and those in other disciplines. He (or she) is a poor general methodologist who has never read Piaget; a poor English methodologist who does not discuss Shaw on spelling or the views of Wordsworth, Coleridge or Matthew Arnold on education in general; a poor Mathematics methodologist who has not read Rignano (24), Hadamard (16) and the opinions of contemporary mathematicians on the teaching of their subject; a poor specialist in psychometry who is unaware of Campbell (5) or Scott-Blair (25) or Weber (29); a poor developmental psychologist who has to confess ignorance of Colette, Alain-Fournier, Turgenev, L. P. Hartley or Carson McCullers; a poor administrator unversed in Mills (21), Wolff (32) and Packard (22); a thinner personality theorist ignorant of Wheelis (30), Pirandello (23) or Camus (6); and the Guidance officer will continue to be the proverbial lightweight without digesting Braithwaite (4).

For the Canadan professional educator this is a necessity, not merely folie de grandeur.

About the second of my two main points, that of the relatively inferior intellectual training provided in education, there was widespread testimony. Members of psychology departments were not noticeably enthusiastic about the quality of work done in Schools of Education; deprecatory remarks were dropped here and there, and one psychologist especially, who had taken when relatively young and foolish a few courses in education, regarded these courses as a special brand of well-organised tomfoolery. Again, as far as research is concerned, there was little interaction between education

and psychology. In one institution the specialist in child psychology complained bitterly to me that he could get no cooperation from specialists in elementary education, despite the fact that he was investigating learning in childhood. One very marked exception to this general rule was the Laboratory of Human Development at Stanford, where students in education and psychology and local teachers worked together on research, apparently with benefit to all.

A third and more urgent testimony came from my scrutiny of the major researches in the west, nearly all of which came from psychology departments. Schools of Education had nothing to compete with Horst's[1] (18) work on differential grade prediction, Mackinnon's on creativity, Festinger's (11) on cognitive dissonance, Guilford's (15) on the structure of abilities and temperament, and Comrey's (8, and many others) on the factorial structure of M.M. P.I. scales. Even the best educational research—that of Kersh (19) in Oregon, Gilbert (13) (14) and Carter (7) on the Berkley campus, Thorpe (27) of the University of Southern California, Keisler at U.C.L.A. and Sheldon (26) at Colorado State College—could not match the intensity of the attack by psychologists. Why, for example, when I visited one institution, did I discover most of the research in education being done by a social psychologist? To recapitulate: the answer seems to be that many professional educators have strayed too far from their basic correlative disciplines in arts and science. In particular some of the educational psychologists seemed unacquainted with contemporary theoretical positions so that they were unable to undertake the extensive hypothetio-deductive researches which would distinguish educational investigation. For example the director of the University of Washington's kindergarten is Bijou, a well-known psychologist, who is using this perfect human laboratory to test the validity of Skinner's learning theory (3).

Doubtless my earnest reader (if he is still with me and has not departed in high rage) will say:

> This is a well-worn theme. We all know by now as a result of Downey's (9a) work at Chicago that Canadian (largely Albertan) professional educators stress intellectual and aesthetic training as the main aim of education while western American educators stress social and patriotic training. Your account is simply a product of cultural shock.

I can reply only that doubtless this is true. I should like to add, however, one further point. Downey discovered that Canadian professional educators chose, as the least important aim of education, training in patriotism and physical skills, whereas the western

[1] In fairness, I should add that Dr. August Dvorak from the School of Education at Washington has worked with Horst on this project. However Dr. Dvorak has done so very largely as a result of his position as a counsellor in the Division of Counseling and Testing.

educator gave training in social and civic skills top priority. This reflects clearly the advantage Canadians enjoy in this respect. The really educated person, as Bertrand Russell was never tired of pointing out, is always critical of the tribal mores, the local as well as the federal variety. Bell (2) has made this point very neatly recently when criticising Lipset's (20) expressed concern that the American intellectual is alienated from the values of his culture.

> An attitude of alienation is one of criticism and skepticism, of acknowledging a relationship while being detached. It is one of the necessary safeguards for intellectuals as a permanent posture, lest they become committed to the winds of ideologies and blow hot and cold with them. (2, pp. 497-498).

In brief if we stress anything other than intellectual training as the main aim of education, we are in danger of creating ideologues whose thinking will be distorted because it is activated by myths and stereotypes, not by evidence. As Bergen Evans (10) has put it:

> We think within the framework of concepts of which we are often unaware. Our most earnest thoughts are sometimes shaped by our absurdest delusions. (10, pp. 267-268).

And the result is never salutary.

Consider, as an illustration of this, a current, intensive and growing dilemma in American education, which is a product of a tacit and almost unconscious acceptance on the part of the professional educator of a Marxis position which was developed by Dewey and handed down by Davis and Havighurst. Dewey, describing the origin of philosophic systems in his *Reconstruction in Philosophy,* argued that the non-empirical philosophies[1] have served the function of justifying the existence of the ruling classes in any society. Any current philosophical and psychological speculations are simplified down to the level of the average person (this is particularly easy since the power élite are in control of the communication media) and, hardened into a myth and expressed in a suitable vocabulary and metaphorical idioms, they create a climate of opinion which favours either directly or indirectly those in power. As I have pointed out elsewhere (1), it is possible to regard the mental testing movement as one of these devices seized by aristocrats and used to preserve the class system, because children from a better socio-economic background get higher scores on tests and consequently, it is argued, the tests must be biased in their favour.

Again it is not very difficult to make plausible by quotation the idea which Freud would have rejected indignantly, that psychoanalysis is another stratagen devised to bolster the class structure. For example E. V. Walter (28) talks about

[1]The Marxist would not exclude empiricism, particularly the contemporary brand, which he regards as a bourgeois attempt to remain aloof from the class struggle. (Comfort, 9).

> . . . an aristocratic conception of life that sees humanity as divided into two species: the primitive formless mass, and the elite, which is evolved, refined and purified from the mass and, presumably restrained by minor controls.

If this is so, then it is easy to show that Freud's doctrine is aristocratic. In his *The Future of an Ilusion* (12), Freud makes the following comment:

> It is just as impossible to do without government of the masses by a minority as it is to dispense with coercion in the work of civilisation, for the masses are lazy and unintelligent, they have no love for instinctual renunciation . . . It is only by the influence of individuals who can set an example . . . that they can be induced to submit to the labours and renunciations on which the existence of culture depends. All is well if these leaders are people of superior insight . . . who have attained the height of mastering their own instinctual wishes.

It is even less difficult to forge a link between Freud and the mental testing movement. The key lies in the phrase 'lazy and unintelligent' in the above quotation. For Freud the extent to which one can be intelligent and achieve depends on the extent to which one can inhibit the drive for libidinal gratification. This dimension of inhibition may be inherited characteristics of the nervous system or it may be largely acquired by the carrot-and-whip system of rewards and punishment so prevalent among the middle class (Davis's socialised anxiety). The main result in either case is that the middle-class child will perform better than the working-class child on the academic criterion so frequently used in aristocratic cultures. In toto, then, the ruling classes have supported a psychology with the sort of cognitive and non-cognitive concepts by which they can justify and perpetuate their own position in society.

Most American professional educators see Britain as the prototype of this aristocratic sort of culture and contemporary British education as privileged and aristocratic. The privileged classes are sitting on top of the lower classes and preventing any sort of social mobility by diehard educational practices such as early selection and homogeneous grouping. Americans seem to be irritated by the determinism and constraint implied in such a conception. American education is seen as correspondingly 'democratic' (the younger men usually avoid this term or use it gingerly), classless and encouraging much upward social mobility. It stresses heterogeneous grouping and encourages as many children as possible to finish high school. (For how long these trends will last is unknown.) It should be noticed that at the basis of the American's conception of his culture and education is a belief in the plasticity of the human organism; this is the conception, nurtured by Davis and Havighurst, that within wide limits and time anybody can learn anything he is adequately taught.

The superstructure of this attitude is now crumbling quickly and

many Americans are coming to see it as just a myth. Whyte (31) and Packard (22) provide convincing evidence that there is in the United States a fairly rigid class structure and that upward social mobility is increasingly difficult. Packard points out that the young American is no longer the most aggressively ambitious of all the world's youth and that he has no desire to found his own firm, be an adventurous enterpreneur or an intellectual pioneer. How does this come about, Packard asks, with the much-vaunted American democratic education? And he goes on to ask why there should apparently be in Britain with its relatively aristocratic education no fewer people at least of indeterminate social class.

Moreover two trends are affecting American education which are even more disquieting, partly because they reinforce each other and partly because they threaten the very basis of the American myth. The first uncomfortable fact is that, as Packard has shown, the contemporary split between the upper and lower classes of the social structure is based to quite a large extent on the possession of a university degree. In fact Packard describes the upper class generally as 'the diploma elite' and insists that this diploma is a necessary qualification for entry into one of the higher-prestige occupations. A second uncomfortable fact for the American educator is that the posts open to youngsters leaving school are more complex and require greater educational preparation than ever before. According to the American census for 1940, 31% of the labour force in the United States were classified as farm workers, labourers or household workers. In 1956 only 17% were in these occupations. In 1958, according to a booklet, *The Labour Market and Employment Security* issued by the U.S. Department of Labour, approximately 11 in 20 non-agricultural job openings were for professional, technical, or managerial posts, whereas only 1 in 20 was for unskilled jobs. Accordingly the unskilled or poorly-skilled person is more likely to be unemployed and a far higher percentage of any one age-group is staying at school beyond the legal leaving age than ever before. In 1900 only one young person in 20 of high-school age was in school: currently the figure is 10 out in 20.

These two uncomfortable facts are all contributing to one thing: the pressure of relatively more people than ever before for university and college education. For example I. J. Quillen, Dean of the School of Education at Stanford University, was reported by the *San Francisco Chronicle* of April 13, 1959, as saying:

> The college is now going down the road the high school followed earlier in the century, when it became the accepted fact that most youngsters would finish high school.

Now obviously this is a test of the merit of the basic American

myth that the human organism is plastic, that with a little better teaching the 'average guy' ought to be able to take a decent university degree. Curiously enough the American legislators and educators have not waited for the results of the test but have simply abandoned the myth—with apologies, of course, but nevertheless the defection is there. For example the President of the University of California is reported in the same issue of the *San Francisio Chronicle* as making the remark:

> I don't want to sound anti-democratic, but we can tighten the standards everywhere so as not to spend money on people who cannot benefit from higher education.

Notice that the President does not say that these students cannot benefit because they don't have the intelligence. Wihout the British aristocratic attitude American educators have no respectable vocabulary dealing with 'intelligence' or 'scholastic potential' as a largely hereditray product which determines the extent of academic achievement. As a result contemporary popular articles such as that John L. Cobbs, "What Do Intelligence Tests Prove?" in *The Saturday Evening Post* for May 9, 1959, are obviously bewildered at such a new-fangled idea. The mere fact that educators are now employing euphemistic synonyms for 'intelligence', words like 'aptitude', 'ability' and 'capacity' (and Cobbs uses these) does not stifle his amazement at the contemporary denial that 'college is for everybody' and that higher education is a right not a privilege.

The contemporary educational philosopher in America has also become slightly more aristocratic than he used to be. For example Hook (17) redefines education to mean "equality of opportunity to achieve through schooling an education commensurate with one's capacities". (17, p. 16). Hook then makes what to British educators must be a fairly alarming educational implication of his definition.

> To say that all children have the same right to an education that will enable them to achieve their growth as human beings is not to say that all children have the same right to the same education independently of their capacities. It does not mean that they have the same right to attend the same schools. (ibid. p. 16).

Hook then attempts to deal with the political implications of such a view:

> No matter how generous our hopes for mankind, to be reasonable they must be compatible with the facts of biological variation. Do the facts of biological variation defeat our ideal of democratic education in the sense that it is futile to expect most students to profit by an education defined by the ends we have previously derived? If they do, then, as Jefferson foresaw, the prospects of our survival as a political democracy are extremely dubious. It seems to me, however, that it is quite reasonable to recognise the facts of biological variation in human capacities and still defend democracy in education. It requires that we distinguish between the *function* of schooling and the *content* of schooling and strive to achieve the same function with different content, where content refers not to ends but to courses and methods and materials of study. (ibid. pp. 16-17).

This distinction is unconvincing. Take the example given by Hook himself:

> A dull child will never be able to read as well or as intelligently as a bright child, and an ordinary child will never be able to play as well as a musically gifted child. But both children can be so educated that each enjoys reading and music. (ibid. p. 19).

This last statement, which is presumably what Hook meant earlier by his phrase, 'to achieve the same function', is implausible. In a large measure the extent of one's pleasure will be determined by the extent of one's capacity. It seems to be more reasonable to accept the fact that the average person has many, often outstanding, skills and abilities which are often more interesting and sometimes more useful than the academic variety.

It is ironic that the position of the aristocrats has been enormously strengthened recently by the advent of Sputnik which has aroused a new interest in the gifted child and in the training of the intellect as the main aim of education. In this American dilemma, then, the main combatants either come from, or are associated with, the U.S.S.R.—Sputnik versus an unwitting Marxist account of mental testing. While it is difficult to forecast the victor, my money is on the American myth which will be almost impossible to dislodge. I recall in this connection the story of Bergen Evans of H. L. Mencken's fabricated tale of the first bathtub installed in America and of how the yarn snowballed as a historical fact despite Mencken's repeated protestations that it was just a joke. So too with the basis of the American myth. I have a notion that it will take more than Sputnik to control it.

And such dilemmas, perhaps a trifle more ignoble and less spectacular, will go on occurring relatively frequently in any area where the primacy of intellectual training as an educational aim is not asserted. The thinking man, the thinking child, is he who refuses to have a settled opinion without evidence, who insists on the right to be sceptical and to doubt. Such a man is and will continue to remain an odd fish. He will meet the opposition of organised groups in society who will allow him to be a technician, but not an intellectual, and who imagine that the larger acts of lunacy are the products of sanity if everyone believes them to be so. The isolated nature of this individual is put very well by Evans.

> Until about a hundred years ago rational men lived like spies in an enemy country. They never walked abroad unless disguised in irony or allegory. To have revealed their true selves would have been fatal.
>
> Today their status is more that of guerillas. They snipe from cover, ambush stragglers, harass retreating rear guards, cut communications, and now and then execute swift forays against detached units of the enemy. But they dare not yet risk an open engagement with the main force; they would be massacred. (10, p. 3).

Despite the fact that these guerillas are small in number, any country which regards intellectual training as the prime function of education will have its due proportion of them and will benefit by their refusal to suspend the questioning spirit in the face of the obiter dicta of the mandarins. I trust that Canada is such a place.

REFERENCES

1. Anderson, C. C. Explanation and the neo-Marxist theory of mental testing. *Alberta Journal of Educational Research.* This issue.
2. Bell, D. Comment on 'American Intellectuals: Their Politics and Status.' *Daedalus,* Summer 1959, 495-498.
3. Bijou, S. Operant or instrumental studies with children using baseline measures. Mimeographed.
4. Braithwaite, E. R. *To Sir, With Love.* London, the Bodley Head, 1959.
5. Campbell, N. R. *Measurement and Calculation.* London: Longmans, Green and Co., 1928.
6. Camus, A. *The Rebel.* New York: Vintage Books, 1958.
7. Carter, Harold D. *California Study Methods Survey.* Los Angeles: California Test Bureau, 1957.
8. Comrey, A. L. A factor analysis of items on the MMPI hypochondriasis scale. *Educ. and Psychol. Measmt.,* 1957, 17, 568-577.
9. Comforth, M. *In Defence of Philosophy.* New York: Lawrence and Wishert, 1954.

9a. Downey, L. The task of the public school as perceived by regional subpublics. Unpublished abstract of a doctoral dissertation, University of Chicago, 1959.

10. Evans, Bergen. *The Natural History of Nonsense.* New York: Knopf, 1957.
11. Festinger, L. The motivating effect of cognitive dissonance. In: Lindzey. G. (ed.), *Assessment of Human Motives.* New York: Rinehart, 1959.
12. Freud, S. *The Future of an Illusion.* New York: Doubleday Anchor Books, 1957.
13. Gilbert, Luther C. Speed of processing visual stimuli and its relation to reading. *J. Educ. Psychol.,* 1959, 50, 8-14.
14. Gilbert, Luther C. Saccadic movements as a factor in visual perception in reading. *J. Educ. Psychol.,* 1959, 50, 15-19.
15. Guilford, J. P. *Personality.* New York: McGraw-Hill, 1959.
16. Hadamard, J. *The Psychology of Invention in the Mathematical Field.* Princeton: Princeton University Press, 1945.
17. Hook, S. The ends and content of education. *Daedalus,* Winter 1959, 7-24.
18. Horst, Paul. Differential prediction of academic success. Mimeographed booklet, University of Washington, May, 1959.
19. Kersh, Bert Y. The adequacy of 'meaning' as an explantion for the superiority of learning by independent discovery. *J. Educ. Psychol.,* 1958, 49, 282-292.
20. Lipset, M. American intellectuals: their politics and status. *Daedalus,* Summer 1959, 460-486.
21. Mills, C. W. *The Power Elite.* New York: Oxford University Press, 1956.
22. Packard, V. *The Status Seekers.* New York: McKay, 1959.
23. Pirandello, L. *Naked Masks.* New York: Dutton, 1958.
24. Rignano, E. *The Psychology of Reasoning.* New York: Harcourt, Brace, 1923.
25. Scott-Blair, G. W. *Measurements of Mind and Matter.* London: Dobson, 1949.

26. Sheldon, M. Stephen. Conditions Affecting the Fakability of Teacher-Selection Inventories. *Educ. and Psychol, Measmt.*, 1959, 19, 207-219.
27. Thorpe, L. P. *Studying Social Relationships in the Classroom.* Chicago: Science Research Associates, 1959.
28. Walter, E. V. The masses and the elite. In: *Dissent.* New York: Grove Press, 69-72.
29. Weber, Max *The Protestant Ethic and the Spirit of Capitalism.* New York, Scribner, 1930.
30. Wheelis, Alan. *The Quest for Identity.* New York: Norton, 1958.
31. Whyte, W. *The Organisation Man.* New York: Simon and Schuster, 1956.
32. Wolff, L. *In Flanders Fields.* New York: The Viking Press, 1959.

PROBLEM SOLVING AND ACHIEVEMENT

Marion D. Jenkinson and Dorothy M. Lampard
Faculty of Education
University of Alberta

Background of the Problem

Twenty years ago the term "problem solving" would have had only one connotation in the classroom—the struggles with problems in arithmetic. In recent years both educational psychologists and teachers have become interested in problem solving methods in general. The former have been primarily concerned with attempting to assess the differentials in problem solving, its functions in thought processes, and its contributions to the learning processes. The latter have been attempting to introduce problem solving techniques into their teaching.

Problem solving has been defined as "the process of overcoming difficulties encountered in the attainment of an objective" (1). Several investigators have probed the ways in which individuals attempt to find solutions to problems (2 and 3). The subjects concerned in these studies have usually been college students or adults. All the research indicates that, although problem solving behaviour varies from situation to situation, from time to time, and from individual to individual, certain basic aspects may be recognized and a sequence of steps in the process has been formulated. More literature is appearing which gives teachers very helpful suggestions of the ways to improve the facility of children in problem solving (4 and 5). However, only one study has as yet been published which indicates that specific training in thinking of a problem solving type will result in improvement in achievement in school subjects (6). High school students were used in this experiment. No study was located which dealt with younger children.

It would appear inevitable that there should be a close connection between the teaching of Enterprise in Alberta and finding solutions to problems, but this may be accomplished without giving specific training in problem solving techniques. Moreover, it is noted frequently in the research concerning drop-outs from the school population, that students who leave prematurely are those who have inefficient study habits. Explicit teaching of problem solving techniques might help to remedy this deficiency. It was therefore decided to undertake a preliminary experiment in the Division II schools of Edmonton to determine whether it was possible to teach grade IV, V, and VI children, methods of solving

problems and whether this training would result in improved achievement.

Purposes of the Study

(1) To ascertain whether the techniques of problem solving could be taught to pupils in grades IV, V and VI.

(2) To determine whether training in problem solving would result in improved achievement in academic subjects. (Reading and social studies were the two subjects chosen).

Subjects

Eighteen classes in grades IV, V and VI in Edmonton schools were chosen to take part in this experiment. Three classes at each grade level were given training in problem solving techniques, while the other nine classes were tested and used as a control. Only the results of those students who completed both the initial and final tests were included in the calculations.

Method

It was decided initially that no one teaching method would be advocated. A variety of methods was evolved for several reasons. This was not only an experimental but also an exploratory study. Though ways of improving the problem solving facilities of children have been advocated, there is little research or empirical evidence which suggests specific teaching methods. The study was planned so that a variety of approaches to the development of problem solving skills were included. Since there is probably no single way to teach a process it was felt that valuable techniques and procedures would be developed if teachers were left free to discover and test possible ways of teaching problem solving.

Before the project began it was discussed with the local superintendents. Nine teachers volunteered to take part in the experiment. These teachers attended seminars which were held at regular intervals. Possible methods of teaching problem solving techniques were discussed and progress reported. After the primary decision to include the sequence of steps suggested by previous research, the participating teachers were encouraged to initiate any feasible method. The sequence of steps introduced as a model was as follows:

(1) Recognizing the Problem
(2) Defining the Problem
(3) Organizing and Recording Observations
(4) Formulating Possible Solutions
(5) Verifying Conclusions
(6) Evaluating Conclusions

Emphasis on techniques and application varied from teacher to teacher and each introduced models, materials and methods as he wished. For instance, one teacher emphasized self-evaluation, whereas another had many workship sessions wherein the pupils discussed their work in groups. Yet another limited the problem solving approach largely to work with the Enterprise, and another linked various steps of problem solving mainly with practice in writing reports and making charts. It was even used by two teachers in Art and Physical Education, two subjects in which the use of problem solving processes is not readily apparent. All teachers attempted to make the pupils aware of the steps involved in the process.

In February 1958, the *S.T.E.P. Reading Test* and the *S.T.E.P. Test of Social Studies* (7), Level IV were given to all pupils in the experimental and control classes. Since reading is a fundamental tool, it was deemed necessary to have some measure of reading achievement of the subjects. The *S.T.E.P. Test of Social Studies* was used because it purports to "measure social studies undertakings, abilities to read and interpret social studies material (map, graph, the printed word, etc.), skill in seeing relationships among basic facts, trends, and concepts, and ability to analyze such material critically" (8). A short reading test was also constructed which gave no guiding question but required the pupil to suggest the solution to an imaginary but feasible problem which he might encounter in real life. The results of this test will be discussed in a subsequent article.

Findings

TABLE I

COMPARISON OF THE INITIAL SCORES AND FINAL SCORES FOR THE READING AND SOCIAL STUDIES TESTS OF THE EXPERIMENTAL AND CONTROL GROUPS

	Exp. Mean	Control Mean	Diff.	t	Sign.	Exp. N	Control N
Reading							
Initial ...	46.69	46.03	.66	.61	NS	224	229
Final	50.19	47.82	2.38	2.14	.05	224	229
Social Studies							
Initial ...	41.63	40.93	.69	.67	NS	224	229
Final	46.63	43.44	3.19	2.99	.01	224	229

From Table I it can be seen that at the beginning of the project, there was no significant difference in the means of the scores on the S.T.E.P. reading test or the S.T.E.P. test of social studies between the experimental and control classes. At the end of the project, however, the pupils in the experimental classes made a mean score

on the reading test of 50.19 as compared to that of 47.82 in the control classes. The difference between these means was significant at the .05 level of confidence. On the final social studies test, the experimental group had a mean score of 46.63 and the control group one of 43.44. This was found to be significant at the .01 level of confidence.

TABLE II

COMPARISON OF THE IMPROVEMENT SHOWN BETWEEN THE INITIAL AND FINAL PERFORMANCES OF THE TOTAL EXPERIMENTAL AND CONTROL GROUPS ON THE READING AND SOCIAL STUDIES TESTS

	READING					SOCIAL STUDIES				
	Initial Mean	Final Mean	Diff.	t	Sign.	Initial Mean	Final Mean	Diff.	t	Sign.
Experimental n 224	46.69	50.19	3.49	3.12	.01	41.63	46.63	5.00	4.76	.001
Control n 229	46.03	47.82	1.78	1.64	NS	40.93	43.39	2.46	2.24	.05

The initial and final scores made on both tests were also analysed to examine the amount of improvement shown by both the experimental and control groups. Table II gives a summary of these results. The difference between the mean scores made by the experimental group on the reading tests given in February and June was 3.49 and this was significant at the .01 level. The control group, however, only made a mean gain of 1.78 during the period of the project, which was not significant. A comparison of the improvement shown between the initial and final results on the social studies tests reveals that the mean differences of both groups were statistically significant. The mean difference of the experimental group, however, was highly significant at the .001 level of confidence.

The results of the analysis by grades of the scores on both the final tests are given in Table III. The only difference in the mean scores made by corresponding grades in both groups which was significant was that made by Grade IV in social studies. But each grade in the experimental group made a higher mean score than did its opposite in the control group, and all these means tended towards significance.

The improvement of each grade, as shown by the initial scores in February and the final ones in June, is presented in Table IV.

The experimental grade IV was the only grade which revealed an improvement in reading that was statistically significant. In social studies, however, the improvement in the means made by each grade in the experimental group was significant. In the control group grade IV also made a significant gain. This was perhaps partially accounted for by one class at this level making outstanding gains.

TABLE III

COMPARISON BY GRADE FOR THE FINAL READING AND SOCIAL STUDIES SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

	Exp. Mean	Control Mean	Diff.	t	Sign.	Exp. N	Control N
Reading							
Grade 6	54.56	52.72	1.84	1.03	NS	72	71
Grade 5	50.63	48.28	2.35	1.25	NS	73	80
Grade 4	45.81	42.88	2.93	1.59	NS	79	78
Social Studies							
Grade 6	51.61	49.51	2.10	1.24	NS	72	71
Grade 5	47.25	44.19	3.06	1.67	NS	73	80
Grade 4	41.53	37.14	4.39	2.73	.01	79	78

TABLE IV

COMPARISON BY GRADES OF THE INITIAL AND FINAL PERFORMANCE OF THE EXPERIMENTAL AND CONTROL GROUPS

	READING					SOCIAL STUDIES				
	Initial Mean	Final Mean	Diff.	t	Sign.	Initial Mean	Final Mean	Diff.	t	Sign.
Experimental										
Grade 6	51.83	54.56	2.72	1.62	NS	47.04	51.61	4.57	2.76	.01
Grade 5	47.44	50.63	3.19	1.77	NS	42.79	47.25	4.45	2.33	.02
Grade 4	41.33	45.81	4.48	2.26	.05	35.63	41.53	5.89	4.04	.001
Control										
Grade 6	50.77	52.71	1.94	1.13	NS	46.63	49.51	2.87	1.65	NS
Grade 5	45.84	48.28	2.44	1.41	NS	43.25	44.19	.94	.59	NS
Grade 4	41.92	42.88	.96	.51	NS	33.37	37.14	3.77	2.21	.05

Summary and Conclusions

Although this was only a pilot study, and the project was under way for a relatively short time, it is evident from these re-

sults that training in problem solving techniques, though not applied to specific subjects, leads to improved performance on academic tests which demand critical thinking and reading. After only five months of such training the pupils in the experimental classes improved significantly over those in the control schools. It is perhaps not surprising that the difference was more pronounced in social studies than in reading, since social studies material probably evokes problem solving techniques more consistently than does that of general reading matter.

It must be emphasised that the study was not primarily designed to assess various methods of teaching problem solving, but rather was intended to ascertain whether the processes could, in fact, be taught successfully at Division II level. Thus no one method was advocated or practised. It is impossible to conclude that any of the methods had greater efficacy than any of the others. Teachers in the experimental schools made the pupils aware of the steps involved in problem solving, and attempted to inculcate the conscious use of these when pupils attacked problems in any subject area. Until recently most of the development of abilities in problem solving has been incidental and unconscious. From this study it would appear that the conscious use of problem solving techniques might be regarded as a learning tool in any area. It is further evident that training in the processes of problem solving can be effective as early as Grades IV, V, and VI.

As in all experiments involving methods of teaching a limitation of this study is the teacher variable. This is partially counteracted by the fact that the methods introduced were not subject centered but were general approaches which were applied to all teaching.

The S.T.E.P. tests are relatively new instruments and some comment on them might be appropriate. In the opinion of the investigators the ceiling of the test appeared to be too low for the grade VI pupils in Edmonton schools. Many of the subjects at this level made nearly perfect scores on the initial tests, and it was thus difficult for them to show any improvement. The next higher level of this series of tests might be more suitable for use in grade VI in Alberta schools. An item analysis of the responses to these tests is being made.

It is hoped that as a result of this pilot study, a larger project will be initiated which will extend over a longer period of time. Other facets which might be examined are the effect of such training on the development of study habits, personality variables and attitudes.

REFERENCES

1. Bingham, Alma. *Improving Children's Facility in Problem Solving.* Teachers' College, Columbia University, 1958, pp. 11.
2. Bloom, B. S. and Broder, Lois J. "Problem-Solving Processes of College Students: An Exploratory Investigation", *Supplementary Educational Monographs,* No. 73. Chicago: University of Chicago Press, 1950.
3. Buswell, Guy T. *Patterns of Thinking in Solving Problems.* Berkeley and Los Angeles: University of California Press, 1956.
4. Bingham, Alma. *op. cit.*
5. Russell, David H. *Children's Thinking.* Boston: Ginn and Co., 1956.
6. Glaser, E. M. "An Experiment in the Development of Critical Thinking", *Teachers' College Contributions to Education,* No. 843, 1951.
7. *Sequential Tests of Educational Progress—Reading—Social Studies.* Cooperative Test Division, Educational Testing Service, Princeton, N.J. Los Angeles 27, California.
8. Sequential Tests of Educational Progress. *Social Studies Manual,* p. 5.

FIFTY YEARS AFTER ALFRED BINET

B. E. Walker*
Faculty of Education
University of Alberta

In the first decade of this century Alfred Binet of Paris gained fame for his pioneer work in testing the intelligence of children. The tradition of an objective, experimental approach to education has been maintained in France and in French-speaking parts of Switzerland and Belgium by a few individuals working for the most part in institutes and laboratories attached to universities. At Geneva, for instance, Claparède studied the emotional development of children and Piaget their intellectual development. Paris, Lyon, Brussels have also been centres of research in education. Unfortuately, French universities have been lagging in their acceptance of the new social sciences, and institutes of education in French-speaking parts of Europe suffer from limitations of staff and resources.

Since World War II government agencies and school administrative systems in France have shown interest in the possibilties of research in education. The Ministry of Education in Paris has expanded the research facilities of the National Pedagogical Institute, creating a Service of Educational Research with Roger Gal as Director. The new interest of the central government is also apparent in the creation of the international pedagogical Institute at Sèvres and the establishment of several *lycées pilotes.* A government bureau of statistics conducts studies relating to student populations at the secondary and higher levels. A research center for for improvement of audio-visual aids is located at St. Cloud.

What are the current problems in French public education requiring study and solution? From interviews with French school administrators and research workers the writer gathered that the most urgent problems are the following: (1) What should be the nature of the school organization and programs for the age group 12-15 years? (2) How can the schools best guide adolescents into appropriate programs? (3) How can the schools integrate science into their programs? In considering all these problems it should be borne in mind that French secondary schools in their classical and modern sections still follow Renaissance traditions in offering a program which is largely literary in nature and designed for an élite.

The educational studies completed to date in French-speaking

*Dr. Walker has just completed a year of study and travel in Europe.

parts of Europe show a preoccupation with instruction in the primary schools. Investigators at Geneva, Brussels, and Paris have cooperated to determine experimentally the optimum grade placement of certain grammatical concepts. Considerable work has been devoted to the improvement of instruction in primary reading. Fernand Hotyat in Belgium has conducted a study designed to improve instruction in mathematics. Jean Piaget has been investigating the ways in which children develop orderly concepts of causality and probability out of the buzzing confusion of impressions and events; he and his co-workers in Switzerland have recently established an international center for genetic epistemology. French government agencies have completed massive studies in the areas of school retardation, the teaching of vernacular language, and the relation of intelligence to social factors.

The teachers in French schools are beginning to make worthwhile contributions in educational research through their individual efforts and through their professional organizations known as *syndicats*. As an example of a voluntary study of a local nature the writer can cite an investigation by two secondary school teachers of the total school population of the town of Aix-en-Provence. The project involved the compilation of six items of data for each of eighteen thousand subjects. Machines were used to tabulate the data and explore relationships among the educational and sociological factors. The investigators drew conclusions and made recommendations which would be helpful to town planners and school administrators.

Ever since Binet helped to direct education into experimental paths, enthusiastic individuals in French-speaking parts of Europe have continued along the same paths, but have found themselves isolated from each other. In order to overcome the disadvantages of isolation, research workers from France, Belgium, and Switzerland have been meeting together informally in recent years. The formal organization of the International Association of Experimental Education (French language) was completed in 1958 and should make possible a more concerted attack on current educational problems. The official organ of the Association is the *Courrier de la Recherche Pédagogique* edited by Roger Gal in the offices of the National Pedagogical Institute at Paris.

REFERENCES

Periodicals—

Bulletin de l'Ecole pratique de psychologie et de pedagogie. Lyon: Université de Lyon.

Courrier de la Recherche Pedagogique. Paris. Institut pédagogique national.

Paedagogica Belgica. Anvers: de Sikkel.

Articles and Books—

Coste, J. P. et Gluck J. *La population scolaire d'Aix-en-Provence.* Aix-en-Provence: La Pensée universitaire, 1948.

Goosens, G. et Roller, S. "Recherches expérimentales sur la connaissance de l'accord du participe passé à Genève et à Bruxelles." *Cahiers de pedagogie experimentale et de psychologie de l'enfant.* No. 11. Neuchâtel: 1954.

Hotyat, F. "Educational Research in French-Speaking countries: Belgium, France, Switzerland. *Review of Educational Research.* Vol. XXVII, No. 1. February 1957.

Institut national d'études démographiques de France. *Le niveau intellectuel d' enfants d'age scolaire.* 2 vols. Paris Presses universitaires de France, 1950-1954.

Ministère français de l'éducation nationale. *Le francais elementaire.* Paris: Centre nationale de documentation pédagogique, 1955

Piaget, J. et Inhelder, B. *De la logique de l'enfant a la logique de l'adolescent.* Paris: Presses universitaires de France. 1955.

Service de la recherche pédagogique. Une enquête sur les retards scolaires. Paris: Institut pédagogique national, 1958.

Unesco. *Research in Education.* Education Abstracts, Vol. 4, No. 2. Paris: Unesco, 1957.

A DESCRIPTIVE SURVEY OF THE EXTRACURRICULAR PROGRAMS OF THE COMPOSITE HIGH SCHOOLS OF ALBERTA

ARTHUR KRATZMAN
Kellogg Fellow, 1957-58
and
A. W. REEVES
Faculty of Education
University of Alberta

During the past four decades extracurricular programs have developed very rapidly in the high schools of North America. In any typical large secondary school of Canada can be found a vast array of clubs, athletic teams, social gatherings, assembly programs, dramatic and musical performances, and journalistic enterprises. This array of activities rivals the formal courses of study in making claims upon the interest and attention of students, and is demanding of the time and energies of teachers.

Another recent development in Canadian educational circles has been the mushrooming across the nation of large composite high schools. These schools, offering a number of course patterns, bring together large groups of students whose abilities and interests differ markedly. The problems encountered in attempting to provide adequate curricular offerings for all composite high school students have been major ones. No less difficult has been the task of administering a meaningful and wholesome extracurricular school life for these students. The primary purpose of this investigation was to gain some insight into the practices and problems in this field of reference—the provision of adequate extracurricular programs for secondary schools with large enrollments and diversified student interests and abilities.

Statement of the Problem

It was the purpose of the study to investigate, through an administrative perspective, the actual nature of extracurricular programs being offered in six large composite high schools of Alberta. The intent was to record the activities offered, the types of organization for the programs, the nature of the supervision of activities, the factors concerning student participation, and the details pertaining to the financing, scheduling, and evaluating of extracurricular programs. Also, an attempt was made to ascertain values ascribed to offerings and to make suggestions which might help to remove obstacles which stand in the way of the realization of these values.

Method of Procedure

Data were obtained from six large composite high schools of Alberta regarding the administration of their extracurricular programs for the 1957-58 school year. Most of the information resulted from a structured interview with the principal of each of the schools, and with individual staff members who had been delegated responsible administrative duties by the principal. Filed documents, such as students' union constitutions, papers outlining the duties of sponsors, regulations for special school "boards", lists of awards, and audited accounts were studied and the information contained in them recorded. Comprehensive questionnaires were forwarded to 153 teacher-sponsors and to 140 student leaders, with returns of 93% and 100% respectively. It should be noted that only those persons who were actually involved with the administration of extracurricular programs in the composite high schools provided information.

Most Pertinent Findings

Based upon the treatment of the above data, the following conclusions appear to be particularly significant.

1. There is a great diversity of structures for the central administrative organization of extracurricular programs in the composite high schools of Alberta. The patterns in effect present extremes—from the role of a principal who is all things to all activities, to the involvement of all staff members in a clearly outlined, but relatively complex administrative structure.

2. Program offerings reveal a great emphasis upon athletic pursuits, especially interscholastic pursuits, where a few students participate actively and the majority are spectators.

3. Staff interests, staff qualifications, and the tradition of continuing with established practices appear to be more important factors in a program initiation than are the interests of the student body.

4. Supervisory practices are as numerous as the composite schools themselves. In general, teachers volunteer their services as sponsors for a variety of reasons; however, 47% of the entire number of instructors in the six schools surveyed make contribution to the supervision of the extracurricular programs. Because of the emphasis on athletic activities, teachers of physical education bear the brunt of supervision. They are responsible for over one-half of the total hours spent on extracurricular supervision in the six schools.

5. Three methods have been used in attempting to weight activities in terms of supervisory load. One school expects that all teachers will share the load; four give scanty remuneration for

sponsorship; three schools give limited time free from instructional duties to a few teachers. Sponsors themselves are in disagreement as to the nature of an adequate distribution of the load, but the largest number favor the load being carried by all regular classroom teachers, with some compensation in their teaching load.

6. Student participation is highest in athletic activities—outnumbering the non-athletic by four to one. However, overall student participation is rather limited, with from one-fourth to one-third of the school population taking part in the extracurricular life of the school. Participation is neither markedly encouraged nor discouraged in any of Alberta's composite high schools.

7. The financing of extracurricular activities in some of the composite high schools of Alberta approaches $4,000 a year, with the greatest amount of money going to the support of interscholastic sports and to the publication of yearbooks. Financial structures for the administration of extracurricular monies are as numerous as the schools. Such structures vary from the principal-treasurer who alone handles the finances, to the student-treasurer who is allowed to manage all financial undertakings save auditing.

8. The great majority of sponsors and extracurricular student leaders attribute much value to the extracurricular programs. Although they see little relationship between academic progress and participation in extracurricular activities, their responses indicate that the latter have inherent values which are not to be found in the day to day activities of the regular program.

9. Evaluation of the extracurricular programs of the composite high schools of Alberta has received little attention. This is inevitable, since five of the six schools surveyed have no formulated goals for the programs to achieve. Objectives, as propounded by school principals when interviewed, vary greatly and, in at least three cases, bear little resemblance to the commonly cited goals in recent literature on this subject.

Recommendations Resulting from the Investigation

The following recommendations are based upon the conclusions immediately preceding, and are outlined within a theoretical framework established from a review of the most recent and pertinent writings in the field:

1. i. Involve a greater number of staff and students in the administration of activities.
   ii. divide the supervisory load more evenly.
   iii. establish clear lines of authority.
   iv. provide staff members with a clear perspective and understanding of the program.

2. The administration of the extracurricular program should give students more opportunities to plan, to make decisions and to accept responsibility. In some schools the provision for such student participation may be found in the constitution but not in practice.

3. Each school should, from time to time, survey current student interest. Unless this action is taken activities tend to be influenced unduly by tradition and by the particular interests and capabilities of staff members. This may well reveal a need to modify the program of extracurricular activities. It might even reduce the attention given to competitive, spectator sports.

4. School administrators should endeavor to secure wider teacher participation in the supervision of the extracurricular program.

5. Sponsors and supervisors should be given relief in the form of a reduced teaching load rather than in the provision of extra remuneration. This is considered to be essential if teachers are expected to fulfill their classroom duties and, at the same time, come to look upon the extracurricular activities as an integral part of the school's program.

6. Measures should be taken to enable as many students as possible to gain experience in leadership.

7. Students should assume more responsibility in the handling of school funds. A reliable system of checks and balances, superimposed on the financial organization, should be sufficient guarantee of effective student management.

Conclusion

Administrators need to take time to analyze exactly what they are attempting to do in the field of extracurricular activities. Objectives must be formulated; planning, coordinating, organizing, directing—all the aspects of sound administration need to be implemented in setting up the structure in which the program will work; and the efforts so expended must periodically be gauged to determine how successfully the aims are being realized. Until a school staff has a clear concept of where it is going it can never determine whether the goal has been reached. The move toward the analysis of values which accrue from extracurricular activities, the formulation of objectives, the evaluation of progress—these seem basic to the administration of extracurricular programs in composite high schools.

REFERENCE

Kratzman, Arthur, "A Descriptive Survey of the Extracurricular Programs of the Composite High Schools of Alberta," M.Ed. Thesis, University of Alberta, Edmonton, Alberta. p. 105.

A COMPARISON OF THE PERFORMANCE ON SELECTED STANDARDIZED TESTS TO THAT ON THE ALBERTA GRADE XII DEPARTMENTAL EXAMINATION OF A SELECT GROUP OF UNIVERSITY OF ALBERTA FRESHMEN

D. B. Black
Faculty of Education
University of Alberta

Yearly, the enrollment of Alberta high schools is increasing. This is due in part to the increase in our school age population. But an increasing proportion of students are staying in high school to complete their Grade XII and subsequently are seeking admission to the University. The final stage of evaluation of performance in high school, upon which admission to the University may be gained, is the Departmental Examinations.

The Grade XII Departmental examinations have been openly criticised for some time, most recently in briefs to the Cameron Royal Commission on Education. It is held that they are expensive to produce, administer and score, their reliability is unknown and they are accused of possessing a number of other undesirable psychometric qualities. As part of its responsibilities, the Alberta Joint Committee to Co-ordinate University and High School Curricula has set up a study subcommittee known as the Matriculation Study Subcommittee, to study these tests. The responsibilities of this subcommittee have been outlined by Evenson and Smith (1957, p. 65). One of the studies performed for the Subcommittee, and that which is reported here, was the comparison of the various Grade XII Departmental examinations with equivalent standardized tests possessing desirable psychometric characteristics and suited to large group administration.

Primarily this study sought to compare performance on the Grade IX Departmentals, on the Grade XII Departmentals, and on selected standardized tests with University freshman average. A second aspect of the study involved the interrelationship among the above three groups of test scores. The standardized tests included were seven College Entrance Examination Board's (CEEB) subject area achievement tests, three tests of academic aptitude, and selected parts of the Cooperative English Test.

Sample

The experimental population of the study consisted of 529 matriculants of the 1956 Grade XII class who entered baccalaureate

programs in the University of Alberta in Edmonton as freshmen in the fall of that year.

This particular class was chosen to make use of the score data resulting from a special testing program conducted in Grade XII classes throughout the Province in May 1956. The design of this larger special testing program, however, has resulted in limitations being placed on the present study which make the sampling design worth noting at this time.

Briefly, it provided that every 1956 Grade XII student would write the School and College Ability Test Level I (SCAT 1), the College Entrance Examination Board (CEEB), Scholastic Aptitude Test (SAT) and two CEEB subject area achievement tests. The CEEB achievement tests were so chosen as to match two of Grade XII Departmental examination subjects being taken by the student concerned. The sampling and assignment of test pairs was made using the Department of Education IBM class registration cards for the 1956 Grade XII class. Distribution of all tests was conducted by the Department. All CEEB tests were scored by Educational Testing Service and the scores reported to the Department. The resulting Grade XII sample writing each CEEB achievement test pair is given in Table I.

TABLE I

SIZE OF CEEB ACHIEVEMENT TEST SAMPLE INCLUDING SAMPLE FOR TWENTY-ONE SUBJECT PAIRS*

	English Composition	Biology	Chemistry	Physics	Interm. Math.	Advanced Math.	French
English Composition	1273	.....	.....	.....	.....	.....	.....
Biology	223	1254	.....	.....	.....	.....	.....
Chemistry	119	209	1267	.....	.....	.....	.....
Physics	206	195	227	1223	.....	.....	.....
Intern. Math.	228	224	219	206	1236	.....	.....
Advanced Math.	149	128	146	150	158	880	.....
French	206	215	218	212	180	143	1203

*After Evenson and Smith (1957)

Consequently, of the 1956 Grade XII matriculants entering the University, in the fall of that year, besides SAT and SCAT scores, approximately half of the freshman class could present one of the twenty-one combinations of two scores from College Board achievement testing as well. Such sampling procedures, then, made it

impossible to compare the full College Board battery with the full Grade XII Departmental requirements presently acceptable for admission to the University.

Freshmen, upon entering the University in the fall, also wrote the American Council on Education Psychological Examination and two parts of the Cooperative English Test: Part A. Mechanics of Expression and the reading comprehension section of Part C: Reading Comprehension.

The variables considered in the study are as follows:

i. Grade IX Departmental Examination scores.
Science (Sci IX), Mathematics (Math IX), Social Studies A S.S. A IX), Social Studies B (S.S. B IX), Literature (Lit IX), Reading (Read IX), and the General Test (G.T. IX). The Grade IX General Test is the Laycock Mental Ability Test.

ii. Grade XII Departmental Examination scores.
English 30 (Eng XII), Social Studies 30 (S.S. XII), Mathematics 30, 31 (Math 30, 31), Physics 30, Chemistry 30, Biology 32, French 30. Also used is Sci XII $\overline{X}$ which is the average of Grade XII Departmental examination scores for Physics 30, Chemistry 30, Biology 32. The student may have written only one or as many as all three of these Departmentals. The same method of averaging was used for Math XII $\overline{X}$ for the two senior Grade XII Mathematics courses and For. Lang. $\overline{X}$ which is the foreign language average. It should be noted in this latter instance that scores from different foreign languages as well as two senior courses in the same field, e.g. French 30, 31, may have been combined. However, in this latter regard, students having French 30 scores only were so identified.

iii. Standardized test scores.
Scholastic Aptitude Test Verbal and Mathematics scores (SAT V and V), School and College Ability Test Verbal and Quantitative scores (SCAT V and Q), American Council on Education Psychological Examination Linguistic and Quantitative scores (ACE L and Q), Cooperative English Test: Test A, Mechanics of Expression (Coop. Mech.) and the reading of Test C, Reading Comprehension (Coop. Read.), and seven CEEB Achievement Tests: English Composition (English), Chemistry, Biology, Physics, Intermediate Mathematics (Inter. Math.), Advanced Mathematics (Ad. Math.) and French.

iv. University Freshmen Average was chosen as the best single criterion with which to compare the effectiveness of the various instruments under examination. It should be point-

Table II

INTERCORRELATIONS OF TWENTY-TWO GRADE IX, XII AND SELECTED PSYCHOMETRIC VARIABLES INCLUDING UNIVERSITY AVERAGE AND CEEB ENGLISH TEST

(N=119)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Sci. IX	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
2. Math. IX	.587	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
3. S.S. (B) IX	.572	.523	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
4. S.S. (A) IX	.340	.295	.559	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
5. Lit. IX	.242	.289	.493	.539	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
6. Read. IX	.326	.305	.546	.550	.566	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
7. G.T. IX	.215	.276	.290	.359	.412	.580	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
8. Coop Read.	.256	.194	.416	.356	.473	.584	.584	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
9. Coop Mech.	.312	.383	.486	.592	.417	.610	.490	.473	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
10. ACE L	.359	.259	.551	.496	.538	.662	.482	.621	.580	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
11. ACE Q	.207	.253	.287	.178	.156	.364	.631	.469	.392	.403	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
12. SCAT Q	.261	.347	.268	.159	.141	.368	.532	.338	.435	.385	.592	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
13. SCAT V	.291	.248	.514	.479	.579	.717	.506	.614	.558	.766	.304	.358	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
14. SAT M	.365	.507	.249	.101	.174	.294	.545	.324	.328	.281	.596	.755	.300	.....	.....	.....	.....	.....	.....	.....	.....	.....
15. SAT V	.265	.234	.489	.364	.566	.649	.489	.656	.446	.642	.277	.261	.782	.270	.....	.....	.....	.....	.....	.....	.....	.....
16. Eng. XII	.317	.246	.505	.465	.581	.613	.376	.611	.575	.652	.206	.198	.674	.154	.701	.....	.....	.....	.....	.....	.....	.....
17. S.S. XII	.386	.354	.614	.347	.389	.400	.246	.400	.332	.591	.286	.245	.512	.181	.452	.479	.....	.....	.....	.....	.....	.....
18. Sci. ($\overline{X}$) XII	.298	.420	.374	.181	.224	.208	.124	.204	.269	.324	.223	.327	.248	.271	.205	.297	.473	.....	.....	.....	.....	.....
19. Math. ($\overline{X}$) XII	.294	.393	.145	.081	.043	.147	.238	.205	.221	.151	.339	.451	.195	.466	.063	.146	.139	.425	.....	.....	.....	.....
20. For. Lang. $\overline{X}$	.237	.383	.405	.343	.250	.237	.213	.198	.360	.290	.188	.219	.239	.179	.056	.287	.328	.373	.240	.....	.....	.....
21. CEEB English	.255	.231	.340	.381	.469	.633	.597	.598	.608	.614	.419	.377	.660	.360	.650	.609	.330	.133	.109	.203	.....	.....
22. Univ. Aver.	.219	.296	.138	.024	.081	.130	.182	.205	.228	.234	.292	.425	.167	.381	.123	.168	.334	.614	.412	.214	.194	.....
Mean	76.0	77.9	78.8	75.1	36.4	75.7	75.9	29.0	60.1	35.6	46.0	41.9	38.8	75.7	77.9	70.2	72.6	69.9	64.8	69.5	502	61.7
Std. Deviation	12.7	12.2	12.8	13.3	6.4	14.7	13.8	8.3	9.4	6.4	8.2	5.5	8.7	19.5	19.4	10.5	9.4	10.0	12.6	10.6	90	8.5

ed out, however, that certain tests are good predictors for certain subjects comprising the University Average whereas other subjects predict less well. Similarly, different tests purporting to measure the some ability predict with varying degrees of accuracy performance in different subjects.

The test score data for each student were punched on IBM cards and the inter-correlations calculated by the University of Washington IBM 650 electronic computer. The particular computer program used was prepared by Dvorak and Wright (1958) and is suited only to the data in which all variables were present for each student.[1]

The result of the computer analysis was seven major intercorrelation matrices for each of the College Board Achievement Tests. A sample of one of these seven tables is presented in Table II.

Inasmuch as Table II and it six sister tables present data far in excess of that needed to achieve the goals of the study, only the pertinent intercorrelations for various Grade IX, Grade XII, College board achievement tests and University Freshman Average are summarized below. Comment will be reserved until after these tables are presented.

TABLE III

INTERCORRELATIONS BETWEEN CEEB ENGLISH AND SELECTED BUT RELATED VARIABLES (N=119)

	1	2	3	4	5	6	7	8
1. Social Studies (B) IX	.....	.....	.....	.....	.....	.....	.....	.....
2. Literature IX	.493	.....	.....	.....	.....	.....	.....	.....
3. Reading IX	.550	.566	.....	.....	.....	.....	.....	.....
4. English 30	.505	.581	.613	.....	.....	.....	.....	.....
5. CEEB English	.340	.467	.633	.609	.....	.....	.....	.....
6. Coop Reading	.416	.473	.584	.611	.589			
7. Coop Mechanics	.486	.417	.610	.575	.608	.473		
8. University Average	.138	.081	.130	.168	.194	.205	.228	.....

TABLE IV

INTERCORRELATIONS BETWEEN CEEB BIOLOGY AND SELECTED BUT RELATED VARIABLES (N=49)

	1	2	3	4	5
1. Science IX	.....	.....	.....	.....	.....
2. Science XII $\overline{X}$	.478	.....	.....	.....	.....
3. Biology 32	.395	.892	.....	.....	.....
4. CEEB Biology	.499	.669	.683	.....	.....
5. University Average	.530	.661	.552	.347	.....

[1]The author wishes to express his sincere appreciation of the help in the conduct of this project to the following: Mr. A. B. Evenson, Miss Angela Budnyk and their staff of the Alberta Department of Education; Mr. Graham Stewart, Mr. Walter Hesse and their staff of the Alberta Department of Municipal Affairs; Mr. A. D. Cairns and his staff of the Registrar's Office, University of Alberta, and Dr. Don Decker, Dr. August Dvorak and Dr. Calvin E. Wright, all of the University of Washington.

Table V
INTERCORRELATIONS BETWEEN CEEB CHEMISTRY AND SELECTED BUT RELATED VARIABLES (N=144)

	1	2	3	4	5
1. Science IX	.....	.....	.....	.....	.....
2. Science XII $\overline{X}$	.400	.....	.....	.....	.....
3. Chemistry 30	.345	.891	.....	.....	.....
4. CEEB Chemitsry	.485	.620	.656	.....	.....
5. University Average	.354	.629	.563	.402	.....

Table VI
INTERCORRELATIONS BETWEEN CEEB PHYSICS AND SELECTED BUT RELATED VARIABLES (N=179)

	1	2	3	4	5
1. Science IX	.....	.....	.....	.....	.....
2. Science XII $\overline{X}$	.465	.....	.....	.....	.....
3. Physics 30	.480	.917	.....	.....	.....
4. CEEB Physics	.502	.596	.682	.....	.....
5. University Average	.410	.642	.568	.400	.....

Table VII
INTERCORRELATIONS BETWEEN CEEB INTERMEDIATE MATHEMATICS AND SELECTED BUT RELATED VARIABLES (N=184)

	1	2	3	4	5
1. Math. IX	.....	.....	.....	.....	.....
2. Math. XII $\overline{X}$	.446	.....	.....	.....	.....
3. Math. 30	.408	.961	.....	.....	.....
4. CEEB Inter. Math.	.461	.700	.715	.....	.....
5. University Average	.226	.380	.300	.281	.....

Table VIII
INTERCORRELATIONS BETWEEN CEEB ADVANCED MATHEMATICS AND SELECTED BUT RELATED VARIABLES (N=246)

	1	2	3	4	5	6
1. Math. IX	.....	.....	.....	.....	.....	.....
2. Math XII $\overline{X}$	.507	.....	.....	.....	.....	.....
3. Math. 30	.474	.869	.....	.....	.....	.....
4. Math. 31	.446	.924	.615	.....	.....	.....
5. CEEB Ad. Math.	.483	.557	.568	.455	.....	.....
6. University Average	.387	.603	.509	.572	408	.....

Table IX
INTERCORRELATIONS BETWEEN CEEB FRENCH AND SELECTED BUT RELATED VARIABLES (N=116)

	1	2	3
1. French 30	.....	.....	.....
2. CEEB French	.671	.....	.....
3. University Average	.470	.354	.....

TABLE X

CORRELATION OF GRADE XII AND CEEB TEST VARIABLES WITH FOUR TESTS OF ACADEMIC APTITUDE

Variable	A.C.E.		SCAT		SAT		Laycock (IX G/T)
	L	Q	V	Q	V	M	
CEEB English	.614	.419	.660	.377	.650	.360	.597
English 30	.652	.206	.674	.198	.701	.154	.376
CEEB Biology	.537	.105	.550	.145	.578	.279	.250
Biology 32	.351	.005	.268	.085	.353	.159	.113
CEEB Chemistry	.337	.252	.336	.421	.290	.522	.231
Chemistry 32	.174	.089	.203	.330	.130	.322	.055
CEEB Physics	.378	.392	.404	.422	.337	.636	.367
Physics 30	.239	.241	.498	.391	.251	.503	.203
CEEB Inter. Math.	.210	.412	.275	.594	.273	.725	.333
Math. 30	.155	.278	.219	.475	.160	.586	.178
CEEB Adv. Math.	.419	.337	.404	.483	.399	.591	.355
Math. 31	.228	.221	.211	.306	.172	.353	.217
CEEB French	.388	.088	.377	.179	.350	.527	.178
French 30	.249	.117	.251	.275	.186	.224	.176

Examination of the data in the foregoing tables would suggest the following findings and observations:

1. Proportionately few students present Biology 32 as part of the science courses for their matriculation standing for admission to the University. For this study, it implies that the bulk of the Science XII averages ($\bar{X}$) were derived from Physics 30 and Chemistry 30 marks. Further, it raises certain questions about the programs and academic calibre of students who do take Biology 32 as part of their Grade XII program. It is interesting to note that for the year concerned, 3,504 students wrote Biology 32 compared with 2,506 for Physics 30 and 4,311 for Chemistry 30. No clue is offered in the University Calendar for the fact that only the Faculties of Agriculture, Dentistry and Engineering do not accept Biology 32 as part of their admission requirements. It

seems apparent (either in the eyes of the students or the high school counsellors), that Biology 32 is not considered to be as valuable as Physics or Chemistry for University admission.

ii. Correlations of the seven CEEB achievement tests with tests of academic aptitude reveal (with one exception), that these correlations are higher than those found between Grade XII Departmental examination scores and the same tests of academic aptitude. The exception is the CEEB English Composition test. Both the Departmental and CEEB English tests are interesting for another reason in that both show generally higher relationships with the three tests of academic aptitude than do the other test scores from either battery.

This finding is of significance since it is the aim of the test battery designer to make the various achievement tests measures of the knowledge of the subject rather than to have them perform as intelligence tests. Conversely, it would be presumptuous to suggest that achievement and capacity for academic learning are not closely related. The implication of this finding is that because of the higher relationships displayed, the inclusion of a test of academic aptitude for admission purposes with the present Grade XII Departmental examinations, either at the time of their administration in June or before the admission of freshman to University, is worthy of consideration.

iii. Correlations between Departmental examinations and their CEEB counterparts range from .455 to .715. In this regard, these correlations tend to be lower than those reported by Evenson and Smith (1958: Table VII, p. 74) for the total Grade XII sample. This could be accounted for by the selective nature of the sample of the Grade XII population that enter the University.

iv. Correlations with University Average of the CEEB and the matching Departmental examinations reveal relationships invariably favoring Departmental examinations with one exception, CEEB English. However, in the latter instance, neither the CEEB nor the Departmental examination correlate highly with University Average nor is the difference between them significant.

This finding raises an interesting hypothetical question. It is generally conceded that the Alberta Grade XII Departmentals are designed as admission tests for the University of Alberta. They are closely related to the curricula of the University since the freshman courses are assumed to be,

in most instances, a continuation from where the high school courses end. Under such circumstances, it is to be anticipated that performance on these tests would relate more closely to University freshman performance than would the CEEB tests which contain only those elements common to most high school programs in the particular field tested. What, then, would be the validity of the Alberta Departmental examination scores in institutions other than the University of Alberta? It is the opinion of the inventigator, lacking any evidence, that they would probably be similar to those found for the CEEB tests at the University of Alberta.

v. The Grade IX Departmental Examinations with one exception, English 30, correlate higher with CEEB tests than they do with the Grade XII Departmentals. This is not surprising in the light of the findings referred to in ii. above and those of Black (1958) who pointed out the high relationship between the Grade IX Departmentals and a test of academic aptitude (SCAT 3). Notwithstanding, the relationship of University freshman average to the Grade IX Departmental scores is sufficiently high to be of value in the educational guidance of students upon entry to Grade X.

Intercorrelations of Various Tests of Academic Aptitude to Each Other and to University Average

Because this analysis in effect is dealing with seven different samples where single correlations are reported, it is to be noted that because only median values are reported, higher and lower values were found, The exact correlation coefficients for each group were determined from the seven major tables illustrated by Table II.

Examination of the above two tables would suggest the following findings:

i. Correlation coefficients between scores on tests of academic aptitude and University Average are moderately high. The reported correlation coefficients and median correlation coefficients, however, differ little from those reported by Evenson and Smith (1958).

ii. It is evident that in terms of relationship to University Average the order of effectiveness in uredicting University Average would place the CEEB, SAT, Cooperative SCAT, Cooperative ACE, and the Grade IX General Test in that order.

Table IX

CORRELATION COEFFICIENTS BETWEEN VARIOUS ACADEMIC APTITUDE TESTS AND UNIVERSITY FRESHMAN AVERAGE FOR SEVEN SAMPLES OF FRESHMAN TAKING VARIOUS CEEB ACHIEVEMENT TESTS

Academic Aptitude Test Scoree	Sample Writing Particular CEEB Achievement Test							Median
	Eng.	Biol.	Chem.	Phy.	Interm. Math.	Adv. Math.	French	
N.	119	49	144	179	184	246	116	
G.T. IX	.182	.234	.250	.227	.197	.237	.248	.234
ACE L	.234	.414	.347	.294	.274	.338	.208	.294
ACE Q	.292	.234	.090	.176	.162	.191	.194	.191
SCAT V	.167	.197	.298	.322	.245	.285	.263	.263
SCAT Q	.425	.315	.262	.301	.145	.253	.307	.301
SAT L	.123	.291	.319	.271	.230	.296	.275	.275
SAT M	.381	.376	.253	.293	.185	.349	.348	.348

Table X

MEDIAN CORRELATION COEFFICIENTS FROM SEVEN INTERCORRELATIONS MATRICES FOR STUDENTS WRITING SELECTED CEEB ACHIEVEMENT TESTS AND FOUR TESTS OF ACADEMIC APTITUDE

Test	1	2	3	4	5	6	7
1. G.T. IX	.....	.....	.....	.....	.....	.....	.....
2. ACE L	.510	.....	.....	.....	.....	.....	.....
3. ACE Q	.502	.400	.....	.....	.....	.....	.....
4. SCAT V	.556	.790	.268	.....	.....	.....	.....
5. SCAT Q	.522	.298	.592	.236	.....	.....	.....
6. SAT V	.548	.703	.256	.801	.227	.....	.....
7. SAT M	.542	292	504	300	738	285	.....

Conclusions

i. The Departmental examinations with the possible exception of English show higher relationships with the single criterion variable, University Freshman Average, than do the College Board's achievement tests. Further, for the sample drawn the degree of relationship shown favoring the Departmental examinations was such that on the basis of the present evidence, substitution of CEEB achievement tests in place of Departmentals for admission purposes does not seem warranted. This, however, should not be viewed as a final conclusion as it was not possible with the sampling design to evaluate each of the CEEB achievement tests in the role of substituting singly or as a battery for the Departmental examinations in the prediction of success in a variety of University freshman subject areas.

ii. The College Board achievement tests and the Grade XII Departmentals do show good intercorrelations for the sample, but these were not sufficiently high to suggest that even within moderate limits these tests are interchangeable.

iii. The CEEB Scholastic Aptitude Test (SAT) was the best academic aptitude admission test studied. The highest correlation coefficients with the criterion University Freshman Average, were found with this test's scores. However, the validity coefficients are typical of similar comparisons reported for other institutions.

REFERENCES

Black, Donald B. "A Study of the Relationships of the Grade IX Principal's Rating to Performance on the Alberta Grade IX Departmental Examinations." *The Alberta Journal of Educational Research*, IV: 4, December 1958, p. 227-236.

Dvorak, August and Calvin E. Wright. "An IBM Type 650 Program for a Symetric Correlation Matrix with Summed Quantities." *Division of Counselling and Testing Services*, University of Washington, Seattle, Washington: August 1958, 9 pp.

Evenson, A. B., and D. E. Smith. "A study of Matriculation in Alberta." *The Alberta Journal of Educational Research*: III: 2, June 1957, p. 62-71.

..... "A Study of Matriculation in Alberta." *The Alberta Journal of Educational Research*: IV 2, June 1958, p. 67-83.

PROMOTION vs. NON-PROMOTION: II. THE EDMONTON STUDY

WALTER H. WORTH
Faculty of Education
University of Alberta

The present organization of the elementary school requires that teachers make promotional decisions at the end of each school year. This decision is usually made with ease and confidence about most pupils. But for those pupils who may be classified as low-achievers, the promotional decision is often difficult and worrisome. Should they be promoted or nonpromoted?

Research done in the United States (8, 20)) suggests that greater benefits accrue to low-achievers from promotion than from non-promotion. Is this true in Canadian schools where educational programs, and the expectations and attitudes of teachers, pupils, and parents, may be different from those in the United States? The present study (21) seeks to answer this question. It also aims at some measure of control over variables not considered in previous research. The findings may be of assistance to school workers in understanding and eventually solving the promotion problem.

Problem

This study[1] sought to compare the effect of promotion and non-promotion on the school achievement and social-personal development of matched groups of third and fourth grade pupils ordinarily categorized as low-achievers in the Edmonton Public Schools.

Decisions were made among the following three kinds of hypotheses with reference to those aspects of school achievement and social-personal development which were observed:

1. There is a difference between promoted and nonpromoted low-achievers, and this difference is in favor of *promoted* low-achievers.
2. There is a difference between promoted and nonpromoted low-achievers, and this difference is in favor of *nonpromoted* low-achievers.
3. There is no difference between promoted and nonpromoted low-achievers.

[1]The writer wishes to acknowledge the contributions of the supervisory and teaching personnel of the Edmonton Elementary Schools, and the Alberta Advisory Committee on Educational Research. Also the advice and assistance of G. M. Dunlop, R. J. C Harper, and R. S. MacArthur of the University of Alberta, S. C. T. Clarke of the Alberta Teachers' Association, and J. H. Shores, R. S. Jones, H. F. Kaiser, Van Miller, W. J. Moore, and W. O. Stanley of the University of Illinois.

Method

Selection and Description of Subjects

Late in the 1955-56 school year preliminary data on all third grade pupils in the school system were obtained through the administration of the following standardized tests:

1. California Achievement Test, Complete Battery (reading, arithmetic, language), Primary Form CC.
2. Gates Advanced Primary Reading Test, Type 1—Word Recognition, Form 1.
3. Gates Advanced Primary Reading Test, Type 2—Paragraph Reading, Form 1.
4. California Short-Form Test of Mental Maturity, Primary, S Form.

The achievement battery and the reading tests were administered by the classroom teachers following a training session with the writer. A trained research team was responsible for the administration of the intelligence test, and for the scoring of all of the tests. All tests were given in the same order, and at similar time intervals, in an effort to establish some degree of uniformity in the testing situation.

Special care was exercised to avoid influencing the promotional decision with respect to any third-grade child at the end of the school year. An attempt was made to disguise the actual purpose of the investigation, and the results of the testing program were withheld from the teachers. These precautions seemed necessary in order to justify the assumption that the groups subsequently studied were drawn from a population representative of typical promoted and nonpromoted pupils.

In the normal course of events, 107 third-grade pupils were nonpromoted on June 30, 1956. The records of these pupils were examined at the beginning of the next school year to determine their suitability and availability for this investigation. It was found necessary to exclude from the experiment forty-one of the pupils who had been nonpromoted. Some had left the school system, others had been transferred to special education classes, still others were "New Canadians" with only a limited knowledge of English, and for others the preliminary test data were inadequate.

The remaining sixty-six nonpromoted pupils were carefully matched, case-for-case, with pupils who had been promoted to the fourth grade. The matching was done on the basis of sex, IQ, chronological age, and total achievement (California). In addition, an effort was made to secure the promoted partner from the school, which in the judgement of local school officials, served a similar occupational and income group as that served by the school

n which the nonpromoted pupil was located. The two groups of pupils thus identified became the subjects of this investigation.

With respect to sex, the two groups were identical. Each contained 20 girls and 46 boys. A similar sex ratio prevailed among those nonpromoted pupils excluded from the study[2]. Table I shows the characteristics of each group with regard to IQ, chronological age, and total achievement prior to the promotional decision.

TABLE I

COMPARISON OF PROMOTED AND NONPROMOTED GROUPS WITH REGARD TO IQ, CHRONOLOGICAL AGE, AND TOTAL ACHIEVEMENT

Characteristics	IQ	Chronological Age in Months	Total Achievement Grade Score
Mean for promoted group	94.53	108.71	3.97
Mean for nonpromoted group	94.68	107.89	3.91
Standard deviation for promoted group	12.24	6.77	.34
Standard deviation for nonpromoted group	12.17	6.77	.34
Correlation between the promoted and nonpromoted groups	+.99	+.70	+.52

Experimental Conditions and Procedures

During the 1956-57 school year the promoted pupils were enrolled in forty-five grade four classrooms in thirty schools, while the nonpromoted pupils were in forty-four grade three classrooms in thirty-one schools. In June, 1957, the achievement of both groups of pupils was again measured by a trained research team, under similar test conditions, utilizing the same instruments as in the pre-testing[3]. Thus it was possible to observe the effect of promotion and nonpromotion on pupil achievement by comparing the groups with regard to gain scores in the twelve achievement areas measured by the California and Gates tests.

A sociometric test[4], designed to secure information regarding frequency of choice as a desirable and undesirable work and play companion, was administered to all pupils in those classrooms in which one or more of the subjects of the investigation were register-

[2]A higher rate of nonpromotion among boys than girls is apparently not uncommon (9).

[3]Test equivalence was thus assured. It was assumed that the effect of practice was minimized by the time interval between testing periods.

[4]Four questions were asked: (1) Name the three people you would most like to work with, (2) Name the three people you would least like to work with, (3) Name the three people you would most like to play with, (4) Name the three people you would least like to play with.

ed during the experimental year. Each subject was rated by his teacher on the following seven personality traits, as defined and used in the Edmonton reporting system (2): emotional control, creativeness, judgment, cooperation, dependability, courtesy, and work habits. A comparison of the pupils on these data was undertaken in an effort to observe the effect of the pomotion factor on social-personal development.

No attempt was made to influence the treatment accorded by school workers to the pupils of either group during the experiment[5]. However, information was secured by means of a questionnaire completed by each pupil's teacher about selected environmental

TABLE II

SIGN TEST COMPARISON OF PROMOTED AND NONPROMOTED GROUPS WITH REGARD TO SELECTED FACTORS RELATED TO EXPERIMENTAL CONDITIONS

Factor	P<NP −	P=NP O	P>NP +	N	Critical Value of $\underline{x}$	Observed Value of $\underline{x}$	§P Value
Teacher continuity from previous year	4	58	4	8	1	4	.36*
Teacher continunity during experimental year	9	49	8	17	4	8	.50*
Years of experience of teacher	30	15	21	51	19	21	.13*
Years of training of teacher	19	21	26	45	16	19	.19*
Health	8	52	6	14	3	6	.40*
Home and family conditions	15	42	9	24	7	9	.15*
Supplementary reading materials	10	44	12	22	6	10	.42*
Supplementary arithmetic materials	2	57	7	9	1	2	.09*
Supplementary language materials	1	62	3	4	0	1	.31*
Supplementary spelling materials	0	56	10	10	1	0	.001†
Small-group instruction	15	44	7	22	6	7	.07*
Individual instruction	17	34	15	32	10	15	.43*
Incentives, rewards and punishments	10	31	25	35	12	10	.01†
Special treatment	6	58	2	8	1	2	.14*
Days absent	35	3	28	63	24	28	.20*
Class size	40	5	21	61	23	21	.01‡

*Decision in favor of hypothesis of no difference.
‡Decision in favor of nonpromoted group.
†Decision in favor of promoted group.
§The P value is the probability, if the null hypothesis is true, of observing a value of $\underline{x}$ less than or equal to the observed value of $\underline{x}$ (4).

[5]The study attempted to observe the effect of the promotion factor under currently prevailing conditions. It was believed that this would strengthen any subsequent recommendations for modifying existing promotion policies and practices.

and instructional factors judged to be influential in determining school progress. The results of a comparison of the groups on these factors is given in Table II.

Statistical tests[6] did not detect differences in teacher experience, training, or continuity; pupil attendance, health, home and family conditions; small-group and individual instruction; provision of supplementary learning materials in reading, arithmetic, and languages; and special treatment (e.g. referral to auxiliary services or special teachers) for the populations of which the two groups were samples. The nonpromoted were, however, enrolled in larger size classes. The teachers of the promoted, on the other hand, provided more supplementary learning materials in spelling, and made greater use of special incentives, rewards and punishments. For a large number of pupils in both groups, there appeared to be few provisions made for adapting instruction to individual needs.

Statistical Treatment and Analysis of Data

A three-decision procedure (5, 7) was used to decide (1) in favour of the promoted group, (2) in favor of the nonpromoted group or (3) in favor of the (null) hypothesis of no difference between the promoted and nonpromoted groups on each of the achievement, social-personal development, and experimental conditions variables observed. This procedure is equivalent to two one-sided tests being carried out simultaneously (3), thereby permitting the detection of directional differences in favor of either group, or a retention of the hypothesis of no difference.

When it seemed reasonable to assume that the differences were normally distributed, a t-test of correlated means requiring no assumption of equality of variance (1, 18) was used to decide among the three hypotheses under consideration. For those data for which an assumption of normality was inappropriate, or for which measurement was difficult but judgment between pairs possible, a sign test (12) was utilized.

The statistical tests were so constructed that the probability of deciding upon a difference in favor of the promoted group when the hypothesis of no difference is true was set at the conventional value of .05. The probability of deciding in favor of the nonpromoted group when the hypothesis of no difference is true was also set at .05. Thus the probability of falsely rejecting the hypothesis of no difference is .10.

Findings

Achievement

In the twelve areas of achievement observed, as shown in Table III, the decision was in favor of the promoted group with regard to

[6]These tests are described in the next section.

gain in reading vocabulary, total reading, and arithmetic fundamentals. In paragraph reading the decision favored the nonpromoted. The decision was in favor of the hypothesis of no difference in reading comprehension, word recognition, arithmetic reasoning, total arithmetic, mechanics of English, spelling, total language, and total achievement. In the eight achievement areas for which the hypothesis of no difference was not rejected, the majority of the differences in the sample means were in the direction of the promoted pupils. The gains in all twelve areas for the pupils of both groups were less than the ten-month gains normally expeced of typical pupils on the standardized tests used.

TABLE III

COMPARISON OF PROMOTED AND NONPROMOTED GROUPS WITH REGARD TO GAINS IN TWELVE ACHIEVEMENT AREAS AS MEASURED BY THE CALIFORNIA AND GATES TESTS EXPRESSED IN GRADE SCORES

Achievement Area	Promoted Group		Non-promoted Group		Difference	Observed Value of t	§P Value
	Mean Gain	S.D. of Gain	Mean Gain	S.D. of Gain			
alifornia:							
Reading vocabulary	.78	.58	.51	.53	+.27	+2.76	.004†
Reading comprehension	.69	.49	.55	.53	+.14	+1.46	.07*
Total reading	.71	.40	.51	.37	+.20	+2.80	.004†
Arithmetic reasoning	.71	.41	.72	.49	—.01	— .10	.46*
Arithmetic fundamentals	.82	.59	.55	.50	+.27	+2.75	.004†
Total Arithmetic	.74	.38	.63	.41	+.11	+1.63	.054*
Mechanics of English	.23	.76	.23	.87	+.00	+ .00	.50*
Spelling	.59	.52	.51	.40	+.08	+1.01	.16*
Total language	.42	.44	.36	.47	+.06	+ .76	.23*
Total achievement	.59	.45	.53	.30	+.06	+ .92	.18*
ates:							
Paragraph reading	.54	.75	.76	.77	—.22	—1.76	.04‡
Word recognition	.55	.65	.67	.46	—.12	—1.26	.11*

(Critical value of t with 65 degrees of freedom ± 1.67)

*Decision in favor of hypothesis of no difference.

‡Decision in favor of nonpromoted group.

†Decision in favor of promoted group.

§The P value is the probability, if the null hypothesis is true, of observing a value of t greater than or equal to (or less than or equal to) the observed value of t depending on the direction of the difference between the means. Its derivation is explained in Wallis and Roberts (19).

Social-Personal Development

When the groups were compared on eleven aspects of social-personal development, as indicated in Table IV, the decision was in favor of the nonpromoted group with regard to desirability as a work companion, and dependability. The decision was in favor of the hypothesis of no difference in emotional control, creativeness, judgment, cooperation, courtesy, work habits, undesirability as a work companion, desirability as a play companion, and undesirability as a play companion. On the nine aspects of social-personal development for which the hypothesis of no difference was not rejected, the bulk of the difference in the sample medians were in the direction of the nonpromoted pupils.

TABLE IV

SIGN TEST COMPARISON OF PROMOTED AND NONPROMOTED GROUPS WITH REGARD TO TEACHERS' RATINGS ON SEVEN PERSONALITY TRAITS AND FREQUENCY OF CHOICE BY CLASSMATES AS DESIRABLE AND UNDESIRABLE WORK AND PLAY COMPANIONS

Trait or Choice	P<NP —	P=NP O	P>NP +	N	Critical Value of $\underline{x}$	Observed Value of $\underline{x}$	P Value
motional control	15	42	9	24	7	9	.15*
reativeness	25	26	15	40	14	15	.08*
udgment	13	39	14	27	8	13	.50*
ooperation	22	27	17	39	13	17	.26*
ependability	27	25	14	41	14	14	.03‡
ourtesy	18	34	14	32	10	14	.30*
ork habits	22	29	15	37	13	15	.16*
esirable work companion	38	8	20	58	22	20	.01‡
ndesirable work companion	24	12	30	54	20	24	.25*
esirable play companion	31	6	29	60	23	29	.45*
ndesirable play companion	29	6	31	60	23	29	.45*

*Decision in favor of hypothesis of no difference.
‡Decision in favor of nonpromoted group.

Influence of Observed Differences in Experimental Conditions

It is possible that these findings may reflect, to some degree, the difference in class size noted earlier. However, research on the effects of class size strongly suggests that class size in itself bears little relationship to pupil achievement and is inconclusive concerning the effect of class size on social-personal development (6, 10, 11, 13, 14). It is also possible that differences in teachers' uses of incentives, rewards, or punishments may have played a part in

influencing the results of the experiment. Psychological research has amply demonstrated the significance of incentives, rewards, and punishments for all aspects of learning (15, 16, 17). However, the kinds of incentives, rewards, and punishments that the teachers reported using, such as progress charts, detentions, and special privileges, may not have been as effective in changing behavior as those described in the psychological literature on motivation. If this were so, it would tend to limit the importance to be attached to this factor in interpreting the findings of this study. Gain in spelling did not appear to be greatly affected by the difference between the groups in the provision of supplementary spelling materials.

Discussion

Normal school progress is commonly considered to be one grade each year. Nonpromotion is usually advocated only when it can be demonstrated to be in the best interest of the child concerned. An argument often used to justify nonpromotion is that it improves school achievement. The findings of this study, however, do not support the view that nonpromotion improves school achievement. Low-achieving pupils who are nonpromoted appear to make no greater, and often less, gain in achievement than they do when promoted.

A possible explanation for the apparent inability of low-achievers to profit academically from nonpromotion may lie in the fact that teachers in this investigation reported using few, if any, supplementary learning materials in an effort to adapt instruction to the individual needs of the pupils concerned. It might therefore be inferred that, in large measure, the nonpromoted pupils repeated the regular third-grade program, while the promoted pupils were exposed to the regular fourth-grade program. If this were the case, and assuming that the repetition of partially-learned material is less stimplating than the challenge of new material, then the repetitive nature of the treatment accorded the nonpromoted may account for these low-achievers not exhibiting greater gains in achievement. The lack of special provisions for the nonpromoted suggests the need for teachers to individualize instruction at the point of error for pupils so classified, in an effort to overcome the boredom and wastefulness that may result from mere repetition of the regular program. Such individualization of instruction may well require the development and utilization of special curricula, methods, and materials. The fact that few teachers reported individualizing instruction for the subjects of this study also suggests that existing programs of teacher education, both of a pre-service and in-service nature, may not be as helpful in preparing teachers for this responsibility as they might be.

Another way to meet the instructional problem posed by the low-achiever could be to develop a form of school organization in which promotion and nonpromotion would not occur. Such a form of school organization would eliminate traditional grade-level lines, and the curriculum rigidity that tends to result, and would substitute a flexible method of pupil classification permitting the continuous progress of pupils from one school term to the next. If this were done, and at the same time provisions made for individualized instruction, it might be possible to bring about greater gains in achievement for low-achievers than those observed in this experiment.

There was some inconsistency in the gain scores in reading. The promoted gained more than the nonpromoted on the California measures. The nonpromoted, on the other hand, gained more than the promoted on the Gates measures. The disparity of these results is difficult to explain. It could be that errors in measurement occurred, or that these tests do measure different aspects of reading ability. It also could be that the Gates tests more closely resemble in content and format the exercise materials commonly used in primary reading programs than do the California tests. If this is the case, then the higher gain scores made by the nonpromoted pupils in paragraph reading and word recognition may reflect, to some degree, the nonpromoted pupils' more recent and frequent practice with similar materials occasioned by their repetition of the third grade. Furthermore, it may be that while tasks not challenging in level and at a low level of expectation lead to a low level of achievement, continued practice with these low-level tasks develops a high level of proficiency with them when proficiency is measured by amount accomplished correctly in a given period of time. If this is so, and assuming that repetition of the third grade reading program and the Gates tests both emphasize reading tasks which are at a relatively low level of challenge and expectation, it may afford another possible explanation for the higher gain scores of the nonpromoted pupils in paragraph reading and word recognition as measured by the Gates tests. Conversely, the greater gain of the promoted pupils in reading vocabulary, reading comprehension, and total reading, as measured by the California tests, may reflect the influence of practice with reading tasks at a relatively higher level of challenge and expectation. Regardless of the reason, the conflicting nature of the gain scores in reading points up the need for school workers to exercise special caution in generalizing about pupil achievement on the basis of performance on a single standardized test.

The findings of this study concerning the effect of promotion and nonpromotion on social-personal development seem to run counter

to those reported in the bulk of previous research (8, 20) on this problem. Nonpromotion does not appear to have an adverse effect on the social-personal development of low-achievers. On the contrary, low-achievers who are nonpromoted tend to be rated as high or higher on personality traits, and to be accorded the same, and sometimes better, sociometric status than those who are promoted. This suggests the possibility that the greater gain in achievement made by the promoted low-achievers may have been at the expense of their social-personal development.

Obviously, however, the promotion factor is not the only influence which was reflected in these results. It may be that as members of larger-size classes the nonpromoted pupils' general behavior was less often drawn to the attention of their teachers and classmates, thereby increasing the likelihood of somewhat favorable, or at least rather neutral reactions. It may also be that the nonpromoted pupils, though they experienced less gain in achievement than did the promoted pupils, were better able to match the academic performance of the other pupils in their classes. If this were so, then this factor may also have led to more favorable reaction from their teachers and classmates. In addition, the extent to which the teachers of the nonpromoted pupils may have deliberately sought social acceptance for these pupils in an effort to aid them in accepting the fact of nonpromotion is not known. Likewise unknown, is the effect of the age differential between the nonpromoted pupils and their classmates. Moreover, since no measure of social-personal development was obtained prior to the promotional decision, it is not known what differences existed in this respect at the beginning of the experiment. It is conceivable, therefore, that the findings with regard to social-personal development were conditioned by earlier differences, and thus do not accurately reflect the effect of the promation factor. And finally, the limitations of the measures of social-personal develpoment which were employed also need to be recognized.

A further explanation for the disparity between the results of this investigation and those conducted earlier may be that the effect of the promotion factor on social-personal development varies with age or grade level. It may also be that this disparity reflects a real difference between Canada and the United States in the attitudes of teachers and pupils toward nonpromotion. These considerations, plus the likelihood of some contamination of the results of this experiment by other factors, serve to emphasize the need for further study of the effect of promtion and nonpromotion on social-personal development.

It is interesting to note that almost seventy per cent of the subjects of this investigation were boys. This was a reflection of the

higher nonpromotion rate among boys than among girls in the population from which the nonpromoted sample was drawn. While each pair of pupils studied was of the same sex, the larger proportion of boys in both groups raise the question, however, as to whether some sort of sex difference may still be mirrored in the results of this experiment. Moreover, the fact that more boys than girls tend to be nonpromoted, as has been shown in this and other studies (9), is in itself a matter which merits further consideration by those concerned with promotional decisions.

In total, this study may be viewed as one bit of evidence that challenges the whole structure of the present elementary school. The findings strongly suggest that, as long as the relatively rigid grade placement of both pupils and content persists, neither promotion nor nonpromotion will adequately meet the educational needs of the low-achiever. It appears that, until school workers develop more flexible instructional and organizational policies and practices, the problems associated with promotion and nonpromotion will be slow to diminish, and the low-achiever, as well as many of his classmates, are unlikely to be provided with the educational program they require.

Conclusions

On the basis of the observed results, it seems reasonable to conclude that:

1. Continued reliance upon nonpromotion in itself to improve school achievement is unwarranted. Low achievers who are nonpromoted appear to make no greater, and often less, gain in achievement than they do when promoted. If the practice of nonpromotion is to continue it must be justified on ground other than improved achievement.

2. Nonpromotion does not seem to have as adverse an effect on social-personal development as previous research might lead one to expect. The social-personal adjustment of low-achievers when they are nonpromoted appears to be as good, if not better, than it is when they are promoted. However, further research is needed to facilitate a more clear-cut interpretation of the differences in the social-personal development of promoted and non-promoted pupils.

Recommendations

The findings and conclusions of this study give rise to four major recommendations concerning the promotion problem:

1. That existing promotion policies and practices be carefully re-examined with a view to clarifying the bases for promotional decisions, and reducing the incidence of nonpromotion for the purpose of improving school achievement.

2. That an attempt be made to ascertain why more boys than girls are nonpromoted so that administrative and curricular policies may be modified to account for this sex difference.

3. That experimentation be undertaken to develop a type of school organization which will permit continuous pupil progress, thereby eliminating many of the problems associated with promotion and nonpromotion.

4. That attention be given to the development of special curricula, methods, and materials designed to facilitate individualized instruction so that the educational needs of the low-achiever may be more effectively met under conditions of promotion and nonpromotion, or in continuous-progress plans wherein promotional decisions are not required.

REFERENCES

(1) Dixon, W. J., and Massey, F. J. *Introduction to Statistical Analysis.* New York: McGraw-Hill, 1957, p. 125.

(2) Edmonton Public Schools. *Directions for Using Report Cards in Divisions I, II, and III and Personality Rating Scales.* Edmonton: 1955, pp. 10-13.

(3) Hodges, J. L., and Lehman, E. L. "Testing the Approximate Validity of Statistical Hypotheses." *Journal of the Royal Statistical Society,* Series B (Methodological), 26 (No. 2, 1954), pp. 261-68.

(4) Harvard University. "Tables of the Cumulative Binomial Probability Distribution." *The Annals of the Computation Laboratory of Harvard University,* 35 (1955). pp. 407-23.

(5) Kaiser, H. F. "Directional Statistical Decisions." Bureau of Educational Research, University of Illinois, Unpublished manuscript, date of publication not yet known.

(6) Kemp, L. C. D. "Environmental and Other Characteristics Determining Attainment in Primary Schools." *British Journal of Educational Psychology,* 25 (June, 1955), pp. 67-77.

(7) Lehman, E. L. "Some Principles of the Theory of Testing Hypotheses." *Annals of Mathematical Statistics,* 21 (March, 1950), pp. 1-26.

(8) National Education Association. *Pupil Failure and Nonpromotion,* Research Memo 2. Washington: Research Division, 1959.

(9) National Education Association. *Pupil Promotion Policies and Rates of Promotion,* Educational Research Service Circular 5. Washington: American Association of School Administrators and Research Division, 1958, p. 4.

(10) Otto, H. J., Condon, M. L., James, E. W., Olson, W., and Weber, R. A. *Class Size Factors in Elementary Schools,* Bureau of Laboratory Schools, Publication 4. Austin: University of Texas Press, 1954.

(11) Otto, H. J., and Von Bergersrode, F. "Class Size." *Encyclopedia of Educational Research,* Revised Edition, 1950, pp. 212-15.

(12) Siegel, S. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956, pp. 68-75.

(13) Spitzer, H. J. "Class Size and Pupil Achievement in Elementary Schools." *Elementary School Journal,* 54 (October, 1954), pp. 82-86.

(14) Steffensen, J. P. "The Relationship Between Class Organization in the Elementary School and Achievement in Certain Subject Matter Areas." Unpublished Doctoral dissertation, University of Illinois, 1958.

(15) Symonds, P. M. "What Education Has to Learn From Psychology." and *Teachers College Record,* 56 (February, 1955), pp. 277-85.

(16) Symonds, P. M. "What Education Has to Learn From Psychology: II. Reward." *Teachers College Record,* 57 (October, 1955), pp. 15-25.

(17) Symonds, P. M. "What Education Has to Learn From Psychology: III. Punishment." *Teachers College Record,* 57 (April, 1956), pp. 449-62.

(18) Walker, H. M., and Lev, J. *Statistical Inference.* New York: Henry Holt, 1953, p. 153.

(19) Wallis, W. A., and Roberts, H. V. *Statistics: A New Approach.* Glencoe: The Free Press, 1956, p. 436.

(20) Worth, W. H. "Promotion vs. Nonpromotion: I. The Earlier Research Evidence." *Alberta Journal of Educational Research,* 5 (June, 1959), pp. 77-86.

(21) Worth, W. H. "The Effect of Promotion and Nonpromotion on Pupil Achievement and Social-Personal Development in the Elementary School." Unpublished Doctoral dissertation, University of Illinois, 1959.

EXPLANATION AND THE NEO-MARXIST THEORY OF MENTAL TESTING

C. C. Anderson
Faculty of Education
University of Alberta

Prologue

The present paper, which is in the form of a popular summary, is designed to encourage students to study metal testing in its larger psychological, philosophical and political contexts because this is one of the indispensable ways which any Faculty of Education must travel if it wishes to stop bringing up the rear in the academic handicap.

Contention

Explanation involves a reduction or simplification of a problem to ultimate and causative conceptions. The casual relationship can be interpreted in two ways which are in effect, if not in theory, mutually incompatible. The first, clearly influenced markedly by Hume, regards the casual relationship as consisting of a set of correlations and conjunctions between the component and measured factors of the phenomena being explained. Causality implies and means predictability, 'existential coupling', in Reichenbach's (30) happy phrase, and the distinguished exponents of this doctrine have included Comte, Myerson, Pearson and Dingle. Pearson (28) termed this view 'descriptive' in its application to science, contrasting it with the second view of casual relationship which he termed 'prescriptive" and dismissed as

> . . . an imaginary aspect of science . . . namely that of deducing some inherent necessity in the routine of perceptions, instead of merely describing the routine of perceptions. (28,p. 116)

In this case the casual relationship, the chain of cause and effect, is characteristised by necessity, manifested in an invariable succession of events and established in its elementary form by a covariant associaiton between any two variables.

This latter conception, while more satisfactory and adequate to most inquiries, nevertheless has one great disadvantage in psychological research; it forces the student to make inferences and hypotheses about what goes on, the 'inner events', between the stimulus (if we can observe any) and the response. The isolation and acceptble description of these 'inner events' is never easy, partly because they must remain hypothetical and are therefore awkward to substantiate and partly because there are so many 'events' to choose

from. The construction of a general system of these 'events' deters all but the tough-minded, largely because such a system has to be conceptualised and described in metaphorical terms which may be quite unacceptable to succeeding generations because of their inherent ambiguity and multiple meanings. Valéry (39) has given a vivid description of the origin of the ambiguities with which certain constructs are so replete.

> You must all of you, at some time or another, have noticed this curious fact — that a given *word,* which may be perfectly obvious when used in the ordinary course of communication . . . has a way of becoming almost magically embarrassing, strangely resistant and quite unmanageable in definition, as soon as you withdraw it from circulation with the object of examining it closely . . . as soon . . . as you try to establish its meaning in isolation from its momentary function. It is almost comic to note the difficulty with which we are confronted when we try to establish the *precise* meaning of a word which, in the ordinary routine of life, we use daily to our complete satisfaction. Take, for instance, the word Time. Caught on the wing, it is perfectly lucid, defined and honest. It fulfills its purpose faithfully, so long as that purpose is to form part of a conventional statement . . . But isolate it, clip its wings, and it turns and takes its revenge. You soon become convinced that the number of its meanings is far in excess of its functions. Formerly it was only a *means,* but now it is an *end,* the object of a terrible philosophical desire. (38, pp. 73-74)

Consequently, in the social sciences, inferences or generalisations tend to be detached and only plausible, not necessary, derivations from the observed. The construction of a generalisation is never a purely passive process, never a mere aggregation of observation statements; it contains superadded implications, 'surplus meaning', in Rommetveit's (31) phrase, not among the observed. As Woodger (42) has recognised,

> . . . universal general statements . . . are theoretical in the sense that they say more than is asserted by the conjunction of the observation records. (42, p. 210)

There are three kinds of defence to be adopted against this conceptual-linguistic problem. The first, primarily positivistic in origin and a laudable attempt to get rid of vagueness and ambiguity in verbal usage, attempts to tie concepts strictly to empirical and experimental data. The positivists take the attitude that certain statements, although they cannot be verified directly, are not thereby declared to be meaningless if they can be shown to be reducible to statements which *can* be directly verified. In this case the indirect and more general statement would have the same meaning as the aggregate of the direct statements. The indirect statement is now not merely inferred from, but is regarded as identical with, the direct statement. In psychology, therefore, it is possible to distinguish between the use of indirect statements which can, and those which cannot, be directly verified by this reduction or retrogressive process. This is the basis of MacCorquodale and Meehl's (26) distinction between 'intervening variables' and 'hypo-

thetical constructs'. The constructs involve "the hypothesisation of an entity, process, or event which is not itself observed." (26, pp. 95-96). In any statement about this construct,

> . . . there occur words . . . which are not explicitly defined by (or reduced to) the empirical relations. (op. cit. p. 96)

By contrast the intervening invariable contains

> . . . in its complete statement for all purposes of theory and prediction, no words which are not defined either explicitly or by reduction sentences in terms of the empirical variables . . . (op. cit., p. 103)

In theory this looks encouraging; in practice it has been disappointing. Reichenbach argues, from a philosophical standpoint, against the view that there can ever be any inferred statements identical with groups of direct statements. A much more formidable difficulty, noted by Tolman (37) and M. H. Marx (24) is the specificity of reference of all variables tied to local situations. Marx asks:

> If concepts of a high degree of generality are essential objectives of scientific theory, how can they be obtained through the use of constructs postulated specifically to refer to a particular set of experimental operations? (24, p. 244)

The experimental solution envisaged by Marx and the quite different linguistic one mooted by George (17) do not obviate the point that at heart all intervening variables with any generality of reference are hypothetical constructs, with all their attendant weaknesses. The danger of having these constructs in psychology is that, being 'free-floating' and relatively ambiguous, they can be (and are) swept into the orbits of a variety of general systems and pansophistic syntheses consisting largely of a set of arguable positions not open to proof or disproof. In the physical sciences such high-level hypotheses present few dangers because they can be tested by deductive experimentation; but, as Hayek (19) has pointed out, the use of the hypothetico-deductive method is not nearly so feasible in the social sciences where it becomes possible to hold diametrically opposite hypotheses to account for the same range of facts.

The present writer has already noted (3) this rather undesirable *impasse* and suggested that some handy 'fictions' might be arranged to make more palatable the hard data of psychometric facts. This is really inadequate sophistry because not infrequently these inferential 'fictions', however hypothetically constructed and sketchily established, have social and educational implications. Then they are no longer intellectual sops, dead contrivances, but powered formative influences of quite astonishing significance for lay but informed opinion, and even for the general populace. A good example is provided by Wiggan (41).

> The second warning of biology to statesmanship is brief and simple: that heredity and not environment is the chief maker of men; that it is essentially the man who in the long run makes his environment, much more than it is the environment which makes the man . . . that he is not, as the glib reformer has it, the helpless victim of the passing education, philosophy, and theories of his time, but that, in the germ-cell, from which every man is born, there are resident those powerful personal forces by which he can rise in well-nigh any environment . . . the social and political import of this warning is that nearly all the happiness and nearly all the misery of the world are due, not to environment, but to heredity, that the differences among men are, in the main, due to differences in the germ cells from which they are born . . . (41, p. 35)

Here is an explanation which links pure psychology (it is an explanation of the 'prescriptive' or inferential variety), mental testing (the inference is made to a large extent from the early results and assumptions of the mental-measurement movement) and education (the explanation has clear educational implications). It is based very largely on the inference that the intelligence test measures inherited scholastic potential. Of course the opposite view, equally an inference, has held court for a least a quarter of a century, principally because of the persistence of the sociologists. Perhaps only a discriminating few will recognize that Marx informed this opinion voiced by Joan Simon (35):

> Consciousness is not an individual possession, closely locked away in each individual psyche. Consciousness arises historically from the necessity of intercourse with other men; it has, therefore, from the beginning been a social product. (35, p. 22)

Here we have a situation where cold psychometric facts are purported to be explained by diametrically opposed inferences which though undesirable, are nevertheless powerful because, interpreted differently by divergent philosophies, they have social and educational applications. The educational implication of the genetic explanation is clearly the stratification of pupils in each grade on a basis of differential intelligence and achievement. It is perhaps no accident that the tripartite system of secondary education in contemporary Britain is a product of psychological opinion educated in the cultural tradition of Darwin, Galton, Spearman and Burt. In a similar fashion the environmentalist explanation of differences is ability as a product largely of differences in training and opportunity for development has encouraged the North American conception (at present under heavy fire) of equivalence and uniformity in educational opportunity and training, and of heterogeneous classes. It is quite pointless to argue with Anastasi and Foley (1) in this connection that both scholastic potential and cultural training interact and are interdependent, because the interaction concept has no direct implications for general educational methology[1]. Potent

[1] And, as Peters (29) has pointed out, methods, techniques and procedures are much more important than aims and values in determining the type and quality of an educational system.

movements in education arise from larger social, political and religious systems which emphasise relative rigidity, consistency, generality and uniqueness (their validity denies the validity of other systems in the field). They are potent because this systematic consistency ties them to general positions and thereby aids the 'quest for certainty'. This is no unusual circumstance. In educational philosophy, idealism and empiricism are still the main contenders, with the neutral positivism, as expounded by Peters (29) a pallid thing by comparison because it affects to dispense with the comfortable generality of metaphysics.

The hereditarian-environmentalist controversy has suddenly flared up again in Britain—(5b) (5c) (26a)— and, as is always the case, it seems clear that a certain number of psychologists, sociologists (the latter are always befogged by a vested interest in this connection) and educators are tempermentally incapable of accepting the hard facts and elegant psychometrics with which Sir Cyril Burt has established the new hereditarian case (5d). It is the present contention that the mental-testing movement cannot afford the luxury of sponsoring, or appearing to sponsor, antinomous general theories of this sort; that there should be a reversion to the legalistic conception of explanation in terms of behavioural products and statistical laws; and that this should be supported by contentious forays in areas where there is no, or may never be any, settled opinion but where the flag of plausible inference and suspended judgement has to be shown. As an example of the necessity of this attitude, the following analysis of the neo-marxist theory of mental testing is presented.

Mental testing is a humble discipline, making only practical claims and rising in response to practical needs. A minor and relatively inexact science, its necessary superstructure is neither extensive nor imposing, being confined to a number of indispensable statistical assumptions. This is perhaps its best and only line of defence and justification. Marxism is, on the contrary, a unified Weltanschauung, a pansophistic synthesis of allied disciplines such as philosophy, sociology, psychology and economics. Although, therefore, its propositions and explanatory concepts are only partial interpretations of the observed, they are pressed to be generally applicable to phenomena in the arts and social sciences, and it is characteristic of the attack on mental testing that it is not isolated; there is neo-Marxist musical theory (6), literary theory (8), and biological theory (21), all being aspects of what the Communists term the 'Battle of the Ideas', by which, through the application of Marxist theory to various aspects of Western cluture, they hope to convince us that our speculative arts and sciences are the product of an idealistic, individualistic, decadent because capitalist, culture.

And it is also characteristic that the specific attack on mental testing is delivered from political, philosophic and psychological standpoints which are integrated in the matrix of the Marxist faith.

For present purposes, the cardinal point in Marxism is that consciousness is largely a social product (psychology is a sort of applied sociology). Men enter into social relationshiph which are based on professional/industrial relationships ('the relations of production') and these in turn determine the political structure and the philosophical and psychological rationalisations that help to buttress that structure. In Marx's well-known text:

> In the social production which men carry on they enter into definite relations that are indispensable and independent of their will; these relations of production correspond to a definite stage of development of the material powers of production. The sum total of these relations of production constitute the economic structure of society — the real foundation, on which rise legal and political superstructures and to which correspond definite forms of social consciousness. The mode of production in material life determines the general character of the social, political and spiritual processes of life. It is not the consciosness of men that determines their existence, but, on the contrary, their social existence determines their consciousness. 24, pp. 11-12)

Farrington, (14) in his account of the origin and development of Greek science, has described faithfully how such a division of labour is correlated with a head-hand vocational dichotomy (psychology) and the mind-body impasse of the philosophic schools. He points out that the origin of civilization depends on crafts and technical skills of all kinds, agriculture, horticulture and metallurgy. These skills were the responsibility of the artisans. However there gradually developed the administrative or ruling classes characterised by a theoretical, abstract and political outlook. They managed men while the men managed matter. Further they were literate and the only literature they allowed to appear concerned the abstract speculative sciences of mathematics and astronomy, which had no political implications. The practical or applied sciences, the productive techniques, were handed down exclusively by oral tradition among the artisans, who were therefore denied expression and their techniques could not therefore progress from a multitude of particulars to real generalisations in the form of scientific laws.

Exactly how this polarisation of labour came about, why, for example the all-important technicians should be depressed in this way, is never made clear but its effects are crucial for the Marxists.

> The technical revolution of the Neolithic Age provided the material basis for the civilisation of the Near East. The revolution also determined the social character of the civilisation that was about to arise. It gradually operated to produce a division in society which had not existed before to any comparable extent. At one pole of society it ranged the workers, at the other administrators — here the peasants, the potters, and the smith; there the king, the priests, the nobles. (14, p. 27)

Throughout Greek philosophy and culminating in Plato, there is a clearly expressed contempt for physics, chemistry and mechanics, and an excessive preference for academic and politically harmless disciplines such as philosophy, mathematics and astronomy, a preference perpetuated by the ruling classes in whose financial and political interests a slave economy was essential and which had to be bolstered by philosophic rationalisations.

There were of course practical experimenters—the Ionian School, Thales, Hippocrates, and Empedocles,—but the major Greek tradition is carried on by Pythagoras, Parmenides and Plato, who tried to split thought from practice with their intense interest in logic and mathematics. For Plato[1], the Marxists have reserved their sharpest barbs; in his philosophic position are prominent all the characteristics of the administrative class to which he belonged. An idealistic metaphysics is merged with an aristocratic social philosophy and a genetic view of the transmission of abilities. He distrusted the evidence of the senses as illusory and secondary to the mysterious intuitive knowledge of 'forms'. He ignored the applied sciences, architecture, medicine, engineering and navigation, as unfit for the minds of all except artisans who had none of the abstract intuition necessary to appreciate such knowledge. They could never approximate to it since they dealt entirely with the sensuous world. The correlative political structure was the division into gold, silver, and bronze classes which each begot children in its own likeness so that only occasionally was there a inter-class transfer. Virtues or functions were assigned to every class—the bronze proletariat were to be temperate and industrious, the silver military group courageous and the gold or governor's class wise—and it was the duty of each class to exercise its proper function and no other.

For the Marxist, Plato, corrupted by the slave society in which he lived, organised his perfect society in this way as a rationalisation. He intellect was warped because he did not realise how thoroughly his mind was saturated with an aristocratic capitalist culture which taught him the mind-body division he reproduced so faithfully in his social philosophy—a perfect example of unconscious sociological pressure conditioning a man to accept constructs for which rationalisations had to be found.

This combination of metaphysical idealism, political conservatism and genetic approach to the transmission of cognitive functions is, however, no accident. The privileged class will always be devising other strategems, both speculative and practical, which will in ef-

[1]My colleague Dr. John Macdonald has pointed out accurately that the more subtle Marxists would admit that a few of Plato's positions were relatively progressive.

fect support the same political structure. One of these stratagems is the metal testing movement which, in the opinion of the neo-Marxists, is deliberately designed to bolster the type of social stratification advocated by Plato. For 'proof' they extract the following quotation from Burt (5), ignoring entirely the fact that it is on an a posteriori speculation and not the rationale behind testing, and also assuming that Burt's is a representative opinion. Burt concludes that

> In brief, the philosophical theory that I should offer would not be very far removed from the assuptions to which I am directly driven by the immediate exigencies of factorial work. Roughly it might be described as a modernisation of the old Phatonic doctrine of . . . 'ideas'. Its main principle would be that reality is best described in terms of 'forms', 'structures', or 'Gestalten' — things analogous to the cognitive wholes that we perceive in our own personal consciousness . . . (5, p. 232)

From the mere mention of 'ideas', 'the personal consciousness' and Burt's well-known support of the hereditary transmission of abilities, the neo-Marxists make the wild inductive leap to the assertion that mental testing is a bourgeois weapon in the class war. Accordingly the political and psych-philosophical criticisms of the mental-testing movement by the neo-Marxist are those which they apply to capitalist society and its speculative products.

Politically the movement is regarded as just another device instituted by capitalists to preserve the class system with all its privileges and injustices. The tests are heavily biased in favour of those who have a better socio-economic background, the sons of capitalists of the middle classes, who therefore enter the best schools and jobs to the exclusion of people from poorer strata of society. The accompanying genetic approach to the origin and development of intelligence and to differential fertility in the social classes is designed to support the sagging class divisions of society by announcing its inevitable perpetuation through the hereditary and immutable mechanism of the genes. Accordingly the intellectual élite largely represent, and certainly reflect faithfully the views of, the social élite, whose undemocratic function is primarily to prove to the workers, in the rationale behind, and the practical application of, tests, that they are inadequately equipped by heredity to assume positions of power, that this state of affairs is unalterable, and that therefore they must be ruled financially and politically by a favoured few. As one editorial (12) has it,

> 'Intelligence testing' is now playing a very important role in preserving class privilege in the schools, in apprenticeship schemes in the factories, and in the Armed Forces. The theories of intelligence which are used to justify the tests, however, are only the theoretical abstractions of the practice of testing — a practice which fulfills a definite need of imperialism, which finds it nesessary to freeze the present class divisions in capitalist society by artificial means. (12, p. 289)

Joan Simon ascribes sinister designs to the makers of mental tests,

> The objective result of testing is . . . to uphold and perpetuate class inequalities and to spread reactionary views of 'human nature' which vitiate educational theory, undermine educational effort and dehumanise the educational process. (35, p. 37)

Brian Simon summarises the neo-Marxist position neatly. "Intelligence as measured by the tests is a class-conditioned attribute." (34, p. 62) Earlier it was "an ideological buttress of class rule and a privileged education system" (34 p. 691).

This bending of the facts about mental testing, its origin and intentions, to fit a socio-political hypothesis is clearly evident. That hypothesis is separated by such a wide gap from the mental testing results that it could not follow from them as a necessary logical consequences, but is merely a description of one kind of arrangement of results which might be reorganized to accommodate a different hypothesis with apparently different political implications—for example the contrary view that the genetically well-equipped individual rising above the hindrances of socio-economic pressures, is of considerable importance at crucial moments in political history. However, there is little fixed, necessary and unambiguous connection between test results and nature-nurture controversy and between that controversy and its political implications, and there is even less ground for making an inductive leap beyond the results to the implications. The danger of this sort of high-level speculation lies in its divorce from the level of observation; we are in the realm of explanations and causal predictions which are in fact only interpretations of, not genuine reversible deductions from, the observed. And the best interpretation of this type is that heredity and environment are interdependent and interact. There is no guarantee that available test results will fit either the genetic or the environmentalist hypothesis.

This can be illustrated by Cattell's (7) inaccurate prophecies inferred from the differential birth-rate in the social classes. Investigations have shown small but significant correlations between I.Q. level and size of family, the larger family having the lower I.Q. average. Cattell, adopting a genetic standpoint, lists among the things we had to look forward to, a decline in mean intelligence for the community, a decline in scholastic standards, an increase in delinquency, and increase in unemployment at the skilled level, an increased conservatism shown in a reversion to rigid religious forms, an increased susceptibility towards propaganda, and a shift from democratic towards bureaucratic forms of government. Few, if any of these things have happened. In fact the 1947 Scottish mental survey has established that a lower average test score for larger families was combined with an overall average improvement in test

score. As yet no high-level hypothesis has plausibly explained this coincidence to the exclusion of any other.

The sensible procedure is to work at the level of observation—statements, and criticism of the neo-Marxis political attack on tests is best dismissed at this level — for example Morris's (27) position.

> What do intelligence tests set out to prove? Simply put, they set out to prove firstly, that only a small minority of children (and hence of adults) are capable of 'logical reasoning'; secondly that mass secondary and higher education is therefore impossible; thirdly that the existing class divisions in education are therefore fundamental and eternal. The corollary of these propositions is the existence of a hereditary elite — the present ruling-class; they along wtih the few workers who manage to jump all the hurdles placed in their way by the class-system of education, are the only people fit for advanced educational opportunities. (27, p. 240)

These assertions can be asnwered point by point. Firstly intelligence tests were designed to provide answers, notably how to discriminate between mentally defective and normal children. They were used also to isolate those capable of the highest forms of 'logical reasoning', but not to prove that only a few were capable. On the whole psychologists are not keen to believe that their procedures are deliberately arranged so that they adduce proof only of notions previously entertained. Secondly the fact that intelligence is to a certain extent inherited does not thereby invalidate mass secondary education, and no accepted tester would indulge in what George Bernard Shaw was fond of labelling 'psychic determinism'. Thirdly the results of testing imply nothing necessarily about the political structure which is certainly a product of much more than cognitive abilities alone. It is true that there is a small but significant correlation between I.Q. and socio-economic level, and that the greatest amount of ascent in social mobility, as Hall and Glass (19) indicate, is made by people with a background of work at high school and in further education, but both of these findings are explicable on either an environmentalist or a genetic position. Again no tester assumes that his work cements the existing political structure. On the contrary, one facet of the intelligent person is that, in being more detached, objective and critical, he is potentially classless, the last sort of person to support the élite of wealth alone. Further Vernon (39) has noted that both Burt and Thomson regarded the group intelligence test as basically democratic, and cites a report of one of Thomson's lectures in the Newcastle Daily Journal for 21 November 1921.

> These psychological tests favour the gifted boy with poor advantages, rather than the rich boy with gifted tutors, and are therefore essentially a democtratic method of selection.

Again all children in the public-schools have to jump hurdles, not merely only those of the artisan as Morris suggests, and Floud and

Halsey (16) have demonstrated that, in grammar-school selection in Britain, the abolition of the intelligence test has meant a significant and consistent decline in percentage of children from lower-class homes who were awarded a grammar-school place. Finally the general idea that intelligence tests aid a class conspiracy becomes moonshine when applied to the American culture where the intellectual is suspect and denigrated.

Of course it is always possible that the tests, particularly the sort that are culturally 'loaded', could be used deliberately to restrict the intake of the sons of artisans into the best and more advanced schools and universities, but this manoeuvre would make no appeal to the majority of responsible people in education.

From the psycho-philosophical standpoint, the neo-Marxists offer two criticisms of the mental-testing movement. The first, criticism of 'behaviouristic or mechanical materialism', is shooting at a straw man set up for target-practice because of its arbitrary linkage of behaviourism with environmentalism. Simon (33) imputes (a typical neo-Marxist procedure) to the behaviourist-environmentalist two assumptions: firstly, that the environment is unchanging, and that therefore the I.Q. is unchanging, at least qualitatively. Both assumptions are inaccurate and unnecessary. The environmentalists in mental testing—Fleming (15), Blackburn (4) and Wellman (40)—believe that the environment and consequently the I.Q. change. Forced to admit this, Simon falls back hurriedly to the further imputation that these same environmentalists assume that the changed environment impinges on an unchanging innate factor of intelligence, and that consequently they cannot explain how the I.Q. changes *qualitatively*. Therefore both geneticists and environmentalists come to the same conclusion:

> . . . that in general 'intelligence' cannot be changed. On the one hand, each individual is endowed with a fixed quantity at birth, on the other the environment is regarded as in general unchanging. Each theory denies the possibility of qualitative change. The 'environmentalists', adopting the philosophical standpoint of mechanical materialism, maintain that if a child is put in a new, improved environment he somehow imbibes 'intelligence' from it. They do not attack the basic question as to *how* the individual is modified . . . (op. cit., p. 691)

In this connection, perhaps the least desirable characteristic of neo-Marxism is its tendency to impute to the mental-testing movement a theoretical superstructure of 'genetic idealism'. In fact mental testing depends on no such assuption and any coincidence between tests and genetic constructs is an interpretation of an experimental finding and not a previous and built-in bias.

Further the mere fact that the I.Q. changes qualitatively does not necessarily imply any qualitative change in intelligence as a psychological function. An equally satisfactory way of explaining

the available evidence is to adduce Hebb's (20) distinction between Intelligence A, or innate hereditary potential theoretically unaffected by environmental changes but only to an approximate extent measured by the standard intelligence test, and Intelligence B, a developmental psychological function capable of being enriched by experience. As matters stand at the moment, the evidence (5d) points to the fact that at least 80% of the differences in intelligence among primary school children should be attributed to hereditary causes and even the most ardent of contemporary environmentalists (26a) is willing—albeit a trifle ungraciously—to lower this figure only to 60%. On the level of argument it is reasonable to assert with Burt that the conception of the brain as a tabula rasa would make it unique in the whole biological world. To conclude, the Marxist stand that the concept of innate potential lan be discounted is mere wishful thinking.

Secondly the neo-Marxists attack the head-hand dichotomous conception of the structure of abilities. McPherson lists three aspects of this separatist position:

> . . . the separation of the 'internal' world of man, his consciousness etc., from his practical activity and social environment . . . The separation of psychological processes from the other processes of the body . . . The separation (particularly in considering such problems as intelligence) of that aspect of knowledge concerned with reflection and understanding, from that concerned with perception. (25, p. 233)

More philosophically Shapiro (32) argues that all testers are 'idealists' who maintain that abilities are "Kantain 'things-in-themselves' " (33, p. 21) 'Idealistic' dichotomies like mind-body sensation-perception, spirit-matter, which are based on the notion that the individual contributes a potent and essential factor to cognitive, perceptive, and conative life, are taken to be false. The sociological vestigial correlates of such dichotomies are to be seen in the class divisions of society in which, roughly, the workers represent the materialistic and 'body' aspect of life and the bourgeosie the 'mind'. The philosophical answer of the 'new' Soviet psychology to this hoary mind-body pseudo-impasse, the notion derived from the venerable Marx-Engels position that mind is matter of a rather superior kind, is philosophically antiquated, and of secondary interest to the psychological answer. This takes the positivistic view that consciousness is largely the product of, and is to be measured in terms of, practical activity in a given situation. This unity of consciousness and activity largely accounts for the measurable changes that take place as behaviour is modified. In McPherson's words:

> We can only be intelligent about things, about situations which are closely related to the real life of the individual. We cannot just be intelligent 'in the abstract' as it were . . . The development of the individual's intelligence is not, as Burt and the other idealist

> psychologies suggest, an 'unfolding' of some predetermined innate quality whose progress is merely speeded up or retarded according to the individual's environment and his knowledge of particular facts, etc., but is seen as a constant development in action . . . From this we must conclude that intelligence of the individual is only fully tested in action, in the real life circumstances of the individual. (op. cit., p. 233)

This notion that the quality of consciousness is revealed only in action is an old Marxist[1] story which might be more useful to the neo-Marxists if it could be proved instead of merely being asserted. However if we accept it as valid for the moment, it has certain implications which can be tested experimentally.

Firstly there is the hypothesis that men tend to be intelligent in one situation but not in others, that, in Simon's words, "the degree of 'intelligence' (if we use the term in the bourgeois sense) *differs with the situation.*" (op. cit., p. 692). The present writer (2) has certainly found that there can be a significant amount of function fluctuation between testing occasions for all measureable traits. However the amount of this fluctuation is not constant—in fact sometimes it does not appear at all to any significant extent—and, although it lowers the reliability coefficient for tests, it does not thereby make traits incapable of measurement and decreases only slightly the predictive power of tests by increasing the number of borderline candidates. While, therefore, each testing situation is different for the individual, that difference is not so great as to invalidate the use of tests.

Another implication of McPherson's version of the neo-Marxist psychology is that whereas in everyday situations an 'emotional' response is involved, no account is taken of it in intelligence tests—an unhappy product of the bourgeois factorial separation of cognitive from non-cognitive functions. Simon mentions this point at some length twice (32) (34), linking it with his view that an individual behaves differently in different situations. "It is precisely in this difference of situation, bringing with it differences in emotional attitudes, that differences in individual behaviour manifest themselves." (33, p. 692)

This is again very theoretical. Much of life is routine and uninfluenced by 'emotions'. Certainly both cognitive and non-cognitive processes function together in responses to test items but not in a way significantly different from the extent of their interaction in a later school or vocational situation. The measured interaction is of roughly the same sort in both situations. Generally, there is no indication that with normal people the affective life has a deleterious effect on cognitive-test responses, and Vernon (39) has shown that pressure of examination at 11+ in Britain does not materially

[1]Kapital 1,3.

increase the percentage of children referred because of instability to child guidance clinics.

A third implication of McPherson's position is that the correlation between intelligence tests and later school performance is an artificial product of the fact that they both deal with abstractions, the latter with them especially in the form of isolated school subjects presented in a formal and abstract way. To this extent and in this way the tests reflect the school—and hence the capitalist social —structure from which they spring. This stand has two parts, the educational and the psychological which are so related that a critical analysis of the former can with profit be generalised to apply to the latter.

Davis (11) as a Western educationist, laments the formalism of Western education.

> What proportion of the basic mental problems met by children . . . in their daily life can be solved by having a large standard vocabulary, or skill in reading, or skill in arithmetical processes? . . . Does one observe in more than one out of twenty public-school classrooms any activities which help children to learn how to reason, to analyse, to invent, or does one observe instead activities of memorising, or learning symbols, or reading or listening to predigested solutions by other peeople, and of paraphrasing ('telling the meaning of') other people's words? Most observers would find the latter. (11, pp. 48-49)

However critical thinking and reasoning is best taught not by studying each problem as specific and with unique characteristics of its own but by using it to form tentative abstractions of general applicability. To carry out this operation it is essential to learn symbolism and to read extensively, and to understand other people's solutions is a necessary preliminary to producing ones own. Fundamental reading and counting skills are not quite so useless as Davis paints them.

Similarly, from the psychological point of view, it is assumed that at the core of all intelligent behaviour in every field is a general activity, the swift and accurate perception of relationships, which is a fundamental and necessary explanation of observed findings. It is pointless to interrupt, Joan Simon does, that intelligence is defined differently by different people. These apparent discrepancies are due to the fact that definitions in terms of product and processes are being confused. The product of intelligence can be variously interpreted as effective execution or the ability to profit by experience, but there is only one process, the power of abstracting relationships.

The rationale behind the intelligence test is that it aims to measure this kind of perceptivity which is later to be amplified in life outside the school. And the high relationship between intelligence-test scores and proficiency in life-tests of increasing com-

plexity cannot be attributed solely to the so-called artefact of formal verbal training taken as the underlying verbal skill of both tests and tasks. Critical and inventive thinking, say at the professional level, must be done in abstractions disengaged sooner or later from the matrix of practical activity, and this sort of thinking can be done only if the necessary conceptual intelligence, measured to some extent by the tests, is available.

However this argument at a level too far removed from observation-statements to damit of decisive proof or disproof. Morris's rather ingenuous version of the psychological attack on tests offers a better target for factual disproof.

> In the light of all this it can easily be understood why the C.P.S.U. declared the whole theory behind intelligence testing as profoundly anti-marxist. The tests try to explain phenomena by lifting them from their social context, isolating them from their connections, in order to examine in 'pure' form. They are based on the artificial isloation of the individual from his real social existence . . . They ignore the vital role of the teacher and of education . . . (op. cit., p. 242).

Against this there are certain replies. Firtsly most testers hold no fixed philosophic position about their art. Few would be willing to speculate as openly as Burt and there is no reason to believe that this speculation is an irrational product of bourgeois conditioning. Secondly no tester would seriously deny the value of educational training. The dichtomy of training versus test selection is an artificial one. Obviously good training is indispensable to child development, but this does not mean that selection can be dispensed with. There is evidence that the Russians practice selection extensively. For example Cole (9) reports that competition is very high for entry to the technical universities, and only the very best brains are selected. Kinghorn (22), for contrast, has described how entry to certain American universities is very unselective and the difficulties ensuing.

> The theory of the State university is based upon an interpretation of the word 'democracy' which, in connection with American education has come to mean that schooling to the highest available level is every man's right. Educators have set themselves to provide opportunities for every person to train themselves to the limits of their capacity, and university teachers, though they believe wholeheartedly in the ideal of higher education for all, realise now that the wholesale application of such a theory does not go far towards solving the practical problems brought about through unlimited college entry. Many of them, in fact, are now prepared to admit, however unwillingly, that this costly experiment has turned out to be false. (22, p. 70)

Thirdly testers do not 'explain phenomena' by isolating tests from their social context. They try to predict future performance by a sampling of abilities in a testable situation. Whether these abilities are sufficiently representative of abilities in the life situation can be established (and has been established—for example by Ter-

man and Oden (36)) by a correlation with these situations. If the tests were valueless they would have no such predictive power. In the same way there is no question of isolating a 'pure' form of a function, but of tapping for a function that will react dominantly in more important situations. The test maker is interested in designing a useful test—whether it is 'pure' or not is altogether a secondary consideration. And if the test is within limits useful in prediction, it is difficult to see how much isolation has taken place unless the 'pure' function is the domintant reagent in the life situation, in which case the neo-Marxist position falls to pieces.

Fourthly, the psychological pattern of the structure of abilities (including the head-hand dichotomy) is an interpretation of test results which is forced upon testers and of practical value. The postulation of a large general factor is not merely an intervening hypothesis to account for the interaction of test results; from it can be deduced certain laws that hold good in practice, e.g. backwardness due to dullness will be found in every aspect of a child's work.[1] These tests are empirically designed and do not rise from any idealistic or separatist rationale. If the general pattern of abilities could not be roughly verified, it would be discarded straight.

On this point it is important to note that most Western testers are willing to modify their views in line with available evidence. For example most accept the fact that environmental and genetic factors affect performance on their tests as they are constructed. And Vernon (39) has declared that, because of the effects of further education, the mental age no longer ceases to increase at 15 or 16 years. Western testers are capable of changing their interpretations to meet changing evidence. Rigidity of interpretation is a characteristic (and a weakness) rather of the neo-Marxist position.

This attempt to justify Western mental testing by an appeal to observation and a plea for low-level hypotheses, will be rejected by the neo-Marxist as anti-Marx and therefore deriving from a pro-capitalist early irrational conditioning. It will be dubbed a positivistic position and positivism, as Cornforth is certain, appeals mainly to the middle-class intellectual who wishes to use science and yet who is unwilling to accept what the Marxists take to be necessary implications of scientific thinking: materialism and support for the working classes. The positivist movement supports, sotto voc, idealism and the 'imperialist bourgeoisie'. The mental

[1] B. Simon's (1955) assertion (**The Common Secondary School.** London: Lawrence and Wishart, pp. 50-51) that psychometrists "have not begun to prove that the term 'intelligence' describes anything significant, that is, anything that exists . . . — in the individual's mind" betrays a lamentable ignorance both of experimental results and of the proper linguistic ustge in a science of behaviour.

tester, therefore, in attempting to justify his vocation by standing on his facts, is not adopting "the philosophy of social neutrality", but "a philosophy of capitalism, propagating the ideas of the imperialist bourgeoisie under the guise of non-partisanship". (10. p. 240)

This neo-Marxist attitude, that all anti-Marx sentiments derive from an early and irrational pro-capitalist conditioning, resembles the equally authoritarian obiter dicta of the Freudians, that attacks on Freud stem from an irrational unconscious aggression against a psychology which is revealing too much about people's behaviour. The best defence the mental tester can offer is to demonstrate the resilience of his position by discarding outworn conceptions in the face of new evidence. B. Simon smirks that the tester can no longer pretend that he measures "heriditary qualities with the ringing assurance of the past". (34, p. 13). This is true, but it is a tribute to the tester. We know that certain types of test item have a cultural bias which is serious, but we also know something which the neo-Marxists blindly refuse to admit: that there is a genetic factor behind the responses to tests. For example a social background of a standard type may produce offspring very heterogeneous in intelligence, an impossibility on any theory of sociological conditioning. Further there is a small but significant correlation between the intelligence-level of children separated early from their parents and father's social class.

This is the only fruitful reply to the neo-Marxists—to base and reshap intelligence-testing theory on observed findings, to justify these findings in terms of predictive efficiency, and, if pressed, to explain them inferentially by intervening hypotheses, the merits of which can be decided by experimental test. In this way the mental-testing movement will retain its status as an honest though relatively inexact science.

Envoi

A colleague, Dr. John Macdonald, who is an expert in Soviet education, has been kind enough to read this article and has offered a number of criticisms, to some of which the present writer takes firm exception. Firstly Dr. Macdonald argues that the Marxist position that 'Freedom is the knowledge of necessity' is not an ignoble one. I wonder if he realises that this position has been used to justify the position of every recent and contemporary dictator. He should read some Camus (6a) and think again. Secondly Dr. Macdonald argues that there are some more 'subtle Marxists' than the ones I have quoted, and that the latter are simply carrying out post hoc justifications of Soviet educational policy which are worthless from the point of view of theoretical Marxism. This may be. All I know is that the Simons in Britain are chatter-

ing away with the same old nonsense they purveyed in the years before 1952. Apparently no 'subtler' educational Marxist has arisen.

But perhaps the most fundamental point on which I take issue with Dr. Macdonald is contained in his following point:

> Furthermore, a Marxist would claim, and in my view with some justification, that your defence of mental testing is too ingenuous. Psychometrists do not necessarily confine themselves to the writing of technical statistical articles. They are sometimes impelled to make social judgements on the basis of the data which they and others have collected. These judgements are frequently of a kind which would throw in doubt the Marxist ideal of a classless democracy, and indeed any democratic ideal at all.

Apparently the only acceptably scientific social judgement a tester is allowed to make is one in support of a classless society. Every other judgement is reactionary and unscientific. (This has obvious affinities with the dictum quoted above that there is no freer man than the slave.) Apparently social judgements throwing in doubt the Marxist ideal of a classless society are a contaminated product of early conditioning.

There are two ways in which this curious dogmatic position might be attacked. Firstly, the fact that evidence might explode some cherished social or political ideal does not in the least invalidate that evidence and is certainly no reasoning for hiding the evidence from the light of day. It may be that Dr. Macdonald wishes to live with a happy illusion but I don't think he should recommend the same narcotic to others. Secondly (if we hold the first criticism in abeyance for a moment), Dr. Macdonald apparently has failed to realise the main point of this piece: that his statement is part of the game of assertion that two can play. It can be asserted for contrast that there never has and never will be a classless society and (with Thomson) metal testing is a democratic process and so on. My contention is that, since psychologists are emotionally incapable of agreeing either about the psychometric facts or about their interpretation in terms of the necessary social judgment to be made on the basis of them, the contestants should stop trading punches and concentrate on determining what tests can do for us.

REFERENCES

1. Anastasi, A. and Foley, J.P. *Differential Psychology.* New York, Macmillan, 1949.
2. Anderson, C. C. *Function Fluctuation.* London, Cambridge University Press, 1958.
3. Anderson, C. C. The latest metaphor in psychology. *Dalhousie Review.* 38:176-188, 1958.
4. Blackburn, J. M. *The Framework of Human Behaviour.* London, Routledge and Kegan Paul, 1947.
5. Burt, Sir C. *The Factors of the Mind.* London, University of London Press, 1940.

5a. Burt, Sir C. General ability and special aptitudes. *Educational Research,* 1950, 1, 3-15.

5b. Burt, Sir C. A note on the theory of intelligence. *Brit. J. Educ .Psychol.,* 1958, 28, 281-288.

5c. Burt, Sir C. The evidence for the concept of intelligence. *Brit. J. Educ. Psychol.,* 1955, 25, 158-177.

5d. Burt, Sir C. and Howard, M. The relative influence of heredity and environment. *Br. J. Statist. Psychol.,* 1957, 10, 99-104.

6. Bush, A. Problems of Soviet musical theory. *The Modern Quarterly.* 5:38-47, 1949-50.

6a. Camus, A. *The Rebel.* New York, Vintage Books, 1958.

7. Cattell, R. B. *The Fight for Our National Intelligence.* London, King, 1937.

8. Caudwell, C. *Illusion and Reality.* London, Lawrence and Wishart, 1946.

9. Cole, B. N. Engineering education in the U.S.S.R. *Universities Quarterly.* 11:373-381, 1957.

10. Cornforth, M. *In Defence of Philosophy.* London, Lawrence and Wishart, 1950.

11. Davis, A. *Social Class Influences Upon Learning.* Cambridge, Harvard University Press, 1950.

12. Editorial. Marism and intelligence tests. *World News and Views.* 30:298-9, 1950.

13. Eells, K. and Davis, A. et. al. *Intelligence and Cultural Differences.* Chicago, University of Chicago Press, 1957.

14. Farrington, B. *Greek Science.* London, Pelican Books, 1949.

15. Fleming, C. M. *Adolescence.* London, Routledge and Kegan Paul, 1949.

16. Floud, J. and Halsey, A. H. Intelligence tests, social class and selection for secondary education. *Br. J. Sociol.* 8:33-39, 1957.

17. George, F. H. Formalisation of language systems for behaviour theory. *Psychol. Rev.* 60:232-40, 1953.

18. Hall, J. R. and Glass, . V. Education and social mobility. In D. V. Glass (ed.) *Social Mobility in Britain.* London, Routledge and Kegan Paul, 1954.

19. Hayek, F. A. Degrees of explanation. *Br. J. Phil. Sc.* 6:209-26, 1955-56.

20. Hebb, D. O. *The Organization of Behaviour...* New York, Wiley, 1949.

21. Huxley, J. *Soviet Genetics and World Science.* London, Chatta and Windus, 1949.

22. Kinghorn, A. M. The American state university today. *Universities Quarterly.* 10:69-79, 1955.

23. Marx, Karl. *A Contribution to the Critique of Political Economy.* Chicago, Kerr, 1918.

24. Marx, M. H. Intervening variables or hypothetical construct? *Psychol. Rev.* 55:95-107, 1951.

25. McPherson, A. The philosophical aspects of intelligence. *The Modern Quarterly.* 4:218-37, 1949.

26. MacCorquodale, K. and Meehl, P. E. On a distinction between intervening variables or hypothetical constructs. *Psychol. Rev.* 55:95-107, 1948.

26a. Maddox, H. Nature-nurture balance sheets. *Brit J. Educ. Psychol.,* 1957, 27, 166-175.

27. Morris, Max. Intelligence testing: a class theory and practice. *Communist Review.* August, 238-243, 1950.

28. Pearson, Karl. *The Grammar of Science.* London, Black, 1900.

29. Peters, R. S. Must an educator have an aim? *The Listener.* 60:931-33, 1958.

30. Reichenbach, H. *Experience and Prediction.* Chicago, University of Chicago Press, 1938.

31. Rommetveit, R. Model construction psychology: a defence of 'surplus meanings' of psychological concepts. *Acta Psychol.* 11:335-345, 1955.
32. Shapiro, M. Some notes on mental testing. *Communist Review.* January, 16-21. 1948.
33. Simon, Brian. The theory and practice of intelligence testing. *Communist Review.* October, 687-695, 1949.
34. Simon, Brian. *Intelligence Testing and the Comprehensive School.* London, Laurence and Wishart, 1954.
35. Simon, Joan. Mental testing. *The Modern Quarterly.* 5:19-37, 1949-50.
36. Terman, L. M. and Oden, M. *The Gifted Child Grows Up.* Stanford, Stanford University Press, 1948
37. Tolman, E. C. Discussion from 'Interrelationships between perception and personality: a symposium.' *J. Pers.* 18:48-50, 1949-50.
38. Valery, Paul. Poetry and abstract thought. In J. L. Hevesi. *Essays in Language and Literature.* pp. 69-111. London, Wingate, 1947.
39. Vernon, P. E. *Secondary School Selection.* London, Methuen, 1957.
40. Wellman, B. L. Growth in intelligence under differing school environments. *J. Exp. Educ.* 3:59-83, 1934.
41. Wiggan, A. E. *The New Decalogue of Science.* London, Dent, 1923.
42. Woodger, J. H. Science without properties. *Br. J. Philos. Sc.* 2:193-216, 1951.

NATIONAL GEOGRAPHIC SOCIETY
LAUNCHES 38th YEAR OF WEEKLY
SCHOOL BULLETIN

For the current school year, the GEOGRAPHIC SCHOOL BULLETIN offers the same high standard of accurate, readable text and lively instructive pictures that have aided millions of educators and students since 1922.

This year, nearly 150 separate articles will give new depth and meaning to world events.

The BULLETIN is published and subsidized by the National Geographic Society. First issue for the 1959-60 school year will be in subscribers' hands on Monday, October 5. Each Monday for 30 weeks will bring a new, timely issue—with the exception of the Christmas and Easter holidays.

The widely acclaimed publication is obtained only by writing the School Service Division, National Geographic Society, Washington 6, D.C. Canadian yearly subscription is $2.25.

Any teacher, librarian, educational worker, student, or parent may subscribe. *Teachers may obtain subscriptions for all members in their classes, if desired, but in such cases the copies must be mailed in bulk to one address.*